Combinatorial Aspects of Symmetries on Groups

An MSc dissertation by Shivani Singh

University of Witwatersrand

Faculty of Science

School of Mathematics

Student number : 1225827

Supervisor : Prof Yuliya Zelenyuk

August 2016

Declaration

I hereby declare that this dissertation is my original work, unless otherwise stated, and that all the sources that I have used have been cited and duly acknowledged by means of a complete bibliography. It is being submitted to the University of the Witwatersrand in Johannesburg as fulfillment of a Master of Science degree and it has not been submitted before in any capacity for any other degree or examination to any other university.

Signature

11 November 2016

Date

Abstract

The symmetries on a group G are the mappings $G \ni x \longrightarrow gx^{-1}g \in G$, where $g \in G$. These symmetries have interesting applications to enumerative combinatorics and to Ramsey theory. The aim of this thesis will be to present some important results in these fields. In particular, we shall enumerate the r-ary symmetric bracelets of length n.

Acknowledgements

I would like to sincerely thank my supervisor Prof Yuliya Zelenyuk for warmly receiving me as a student. This dissertation could not have been achieved without her expertise and guidance. Furthermore, I express my gratitude to the National Research Foundation for providing me with the financial support to make this research possible. Lastly and most importantly, I am deeply grateful to my parents who were unwavering in their love, support and encouragement of my studies.

List of Symbols

Sets

$[n]$	$\{1, 2, \cdots, n\}$
$\cap$	Intersection
$\cup$	Union
$\emptyset$	Empty set
$\in$	Element of
$\mathbb{N}$	Natural numbers
$\mathbb{R}$	Real numbers
$\mathbb{Z}^+$	Set of all positive integers
$\mathbb{Z}$	Integers
$\neq$	Not equal to
$\subseteq$	Subset
k-set	Set containing k elements

$Y - \{y\}$	Set Y excluding the point $\{y\}$
Z_1	Set of all integers ≥ 1
$[S]^k$	Set of all k-sets contained in set S

Conventions

$\exists$	There exists
$\forall$	For all
$\Longrightarrow$	Implies
min $\{a, b\}$	Minimum value of $\{a, b\}$
"$\Longleftarrow$"	Backward statement
"$\Longleftrightarrow$"	If and only if
"$\Longrightarrow$"	Forward statement
w.l.o.g	Without loss of generality
w.r.t	With respect to

Group theory

$[G : H]$	Index of H in G
$\mathbb{Z}_n$	Group of all integers modulo n
$\mathbf{e}$	Identity element of G
$\{\mathbf{e}\}$	Trivial subgroup

A_n	Alternating group
$Aut(G)$	Automorphism group of G
$C_g(G)$	Centralizer of an element g in G
C_n, $\mathbb{Z}_n$	Cyclic group of order n
D_n	Dihedral group of order $2n$
$Dih(A)$	Generalised Dihedral group of A
G	Group
G/H	Quotient group of H in G
$G[2]$, $B(G)$	Boolean group of G
gH	Left coset of H in G.
$H \cong G$	H is isomorphic to G
$H \leqslant G$	H is a subgroup of G
$H \ltimes N$	Semidirect product of H and N
$H \times N$	Direct product of H and N
$H \lhd G$	H is a normal subgroup of G
Imf	Image of function f
$Kerf$	Kernel of function f
$N_H(G)$	Normalizer of subgroup H in G

$o(g)$	Order of an element g in G
$Orb(s)$	Orbit of element s
$Stab(s)$	Stabiliser of element s
V_4	Klein four group

Combinatorics

α, β	Incidence functions
δ	Kronecker delta
$\leq_R$	Partial order relation
μ	The Möbius function
π, σ	Partitions
Π_n	Set of all partitions of $[n]$
$\sim$	Equivalence relation
ζ	Zeta function
$I(P)$	Incidence algebra
$L(G)$	Lattice of subgroups of a group G
P	Poset
$\binom{m}{k}$	m choose k

Ramsey theory

$[n]^k$	Set of all k-tuples on $[n]$
$\chi(H)$	Chromatic number of hypergraph H
$\mathfrak{L}$	System of linear homogeneous equations
$G = (V, E)$	Graph with V vertices and E edges
$H = (V, E)$	Hypergraph with V vertices and E edges
$K_{\mathbb{N}}$	Complete graph of $\mathbb{N}$ vertices
K_n	Complete graph of n vertices
L	Combinatorial line
l	Variable word
$l(i)$	Word
$R(a, b)$	Ramsey number
$S(r)$	Schur number
$W(k, r)$	Classical van der Waerden number
$W(k_1, k_2, \cdots, k_r, r)$	Generalised van der Waerden number
AP	Arithmetic Progression

Symmetric colorings

$[\chi]$	Orbit of χ
χ	Coloring
$\leq$	Refinement order relation
ϕ	Euler function
Π_G	Set of all optimal partitions of a group G
A	Finite abelian group
$A\chi$	A-orbit of χ (r-ary necklace)
$B_r(n)$	Number of all r-ary bracelets of $\mathbb{Z}_n$
$B_r^*(n)$	Number of symmetric r-ary bracelets of $\mathbb{Z}_n$
f_g	Symmetry on group G
$G\chi$	G-orbit of χ (r-ary bracelet)
n-gon	Regular polygon with n sides
$N_r(G)$	Number of all r-ary necklaces of G
$N_r^*(n)$	Number of symmetric r-ary necklaces of $\mathbb{Z}_n$
r^G	Set of all r-colorings of G
$S_r(G)$	Set of all symmetric r-colorings of G
$S_r(G)/\sim$	Set of all symmetric r-ary necklaces of G

$St(\chi)$	Stabilizer of χ

List of Tables

List of Figures

Contents

Chapter 1

Introduction

Combinatorics is a branch of discrete mathematics that can be superficially described as the mathematics of counting. This is an acutely insightful field hence many mathematicians, in the past, have found it appealing solely for its aesthetic and recreational value, however today it has a diverse array of applications in many scientific disciplines. Combinatorial results and techniques extend into almost every corner of mathematics. The main objectives of this dissertation is to introduce segments of this field which includes several essential results in Ramsey theory.

The English mathematician called Frank Plumpton Ramsey, in 1928, released the seminal paper [31] and unknowingly spawned a new field of mathematics called Ramsey theory, named entirely in his honor. Unfortunately, Ramsey passed away at the young age of 26 before witnessing the impact of his work[1]. A detailed survey of the fascinating history behind Ramsey theory can be

[1]His paper was published posthumously in 1930.

found in [37]. Loosely defined, Ramsey theory is a subsidiary of combinatorics that is concerned with structures that are preserved under partitions. The classical theorems in this field are noted for their elegance therefore it is vital for a budding mathematician to study them in order to procure a certain level of mathematical sophistication.

Ramsey theory can be summarised into a single emphatic question: can one always find order in chaos? This sounds like a philosophical question so let us illustrate this question with the following well-known example called the "Party problem". Suppose we are required to invite some guests to a party and we wish to find the smallest number of guests to be invited such that no less than m guests will already be acquainted or no less than n guests will not be acquainted. The solutions to m and n are known as *Ramsey numbers*.

In 1927, the result known as *van der Waerden's theorem* was officially published by a Dutch mathematician called Bartel Leendert van der Waerden, in [41] and is another fundamental theorem of Ramsey theory. Van der Waerden's theorem contemplates the colorings of finite sequences and is a direct consequence of the *Hales-Jewett theorem*, a more powerful result that looks at the colorings of arbitrarily large finite dimensional cubes.

A *van der Waerden number* is the smallest positive integer $W = W(k, r)$ such that for all $n \geq W$, every r-coloring of $[n]$ contains a monochromatic k-term arithmetic progression. Intuitively, this theorem can be explained as follows; suppose we partition the set of all positive integers, $\mathbb{Z}^+$, into exactly

two collections, then at least one of these collections must possess arithmetic progressions of arbitrary lengths. One way of visualising partitioning a set into r classes is to think of coloring all the elements of the set with r colors, therefore a 2-coloring of the positive integers will yield monochromatic arithmetic progressions of arbitrary lengths in color 1 or in color 2. This dissertation aims to present some of the traditional theorems in Ramsey theory such as Ramsey's theorem, van der Waerden's theorem, Schur's theorem, Rado's theorem and the Hales-Jewett theorem, among others. For further study, readers are encouraged to consult [16], [17] and [25].

The theory of color symmetry or symmetric colorings originated as a subject in the 1950's where it contributed mightily to the development of crystallography. The central ideas of symmetric colorings are simple yet powerful and has generated many open problems. This dissertation overviews some of the key concepts of color symmetry.

In particular, this dissertation discusses the colorings of regular n-sided polygons (n-gons). Suppose we are given a regular n-gon whose vertices are colored in finitely many colors, say r, in other words each vertex is assigned one of r colors. We can quite clearly see that there are r^n colorings of this n-gon. Two colorings, say χ_1 and χ_2, of a given n-gon are called *equivalent* if χ_2 can be obtained from χ_1 by rotating about the n-gon's center. This relation is an equivalence relation and each of these equivalence classes is called a *necklace*. We use *Burnside's lemma* to enumerate the equivalence classes of r-colorings (necklaces).

Burnside's lemma is a combinatorial result in group theory that has a rather tumultuous history. The lemma was officially stated by Cauchy in 1845, but William Burnside quoted it in 1897, in his book [11]. Burnside apparently attributed it to Frobenius and thought it to be a well-known result, therefore it is also known as the Cauchy-Frobenius lemma. Burnside's lemma supplies an elegant formula for counting mathematical objects when their symmetry needs to be taken into account.

An r-coloring is symmetric if it remains unchanged under some mirror symmetry, whose axis passes through the center of a regular n-gon and through one of its vertices. Formally, a symmetry on a group G is a mapping $G \ni x \longrightarrow gx^{-1}g \in G$, where $g \in G$. An r-coloring of a group G is any mapping $\chi : G \longrightarrow [r]$. An r-coloring is *symmetric* if $\chi(gx^{-1}g) = \chi(x)$ for some $g \in G$.

Counting the number of symmetric colorings and symmetric bracelets poses a more difficult question that this dissertation attempts to answer. The approach or method used to compute the number of symmetric colorings is based on creating the poset of optimal partitions, which is found in [46]. We will use [4] together with [38] to address the combinatorial theory behind this method and the resulting material is modified as well as expanded upon from the articles [43], [45], [46] and [47].

An r-ary bracelet of length n is an equivalence class of r-colorings of vertices

of a regular n-gon, which considers all the rotations and mirror symmetries to be equivalent. In particular, we will derive a formula that enumerates the symmetric r-ary bracelets of length n, found in [44], before proceeding to the conclusion of this dissertation which will provide an assessment of recent breakthroughs and open questions relating to this field.

Chapter 2

Group Theory

This chapter serves as a self-contained introduction to the group theoretical framework required, which consists of several standard results from group theory. This chapter is admittedly not exhaustive therefore several proofs are omitted since they can be found in standard texts such as [14], [23] and [33]. Intuitively speaking, a group is an abstract mathematical object that describes symmetry. A more formal definition is given below.

Definition 2.0.1. Let G be a nonempty set and consider the binary operation $*$, then $(G, *)$ is called a *group* if it satisfies the following conditions.

(1) $\forall g, g_1, g_2 \in G$, $(g * g_1) * g_2 = g * (g_1 * g_2)$, i.e. the operation $*$ is *associative*.

(2) There exists an element $\mathbf{e}$ in G, called the *identity element* of G, such that $g * \mathbf{e} = \mathbf{e} * g$, for all $g \in G$.

(3) For each $g \in G$, there exists an element $\mathbf{e} \neq g_1 \in G$ such that $g * g_1 = \mathbf{e} = g_1 * g$. Then g_1 is called the *inverse* of g in G, $g_1 = g^{-1}$.

We use G to denote the group $(G, *)$ in general. It is fairly trivial to show that the identity element and the inverse of each element in G are unique, see [32]. If the operation $*$ is also *commutative* i.e. if $g * h = h * g$, for all $g, h \in G$, then we call G an *abelian* or commutative group.

Proposition 2.0.1. *Define $a^0 = \mathbf{e}$, then for all positive integers n;*

(1) $a^n = (a^{-1})^{-n}$ *and*

(2) $(ab)^{-1} = b^{-1}a^{-1}$.

Proof. See [23]. □

The cardinality of the group G is called the *order* of G, we use $|G|$ to denote the order of G and if G is a finite group then it will have only a finite number of elements. The order of a single element $\mathbf{e} \neq g \in G$, is the smallest positive integer n such that $g^n = \mathbf{e}$, similarly we denote the order of g by $o(g)$.

Definition 2.0.2. A nonempty subset H of G is called a *(proper) subgroup* of G if it is a group in itself i.e. if it satisfies the three group axioms under the same binary operation that defines G. H is a subgroup of G is denoted by $H \leqslant G$.

Clearly G and $\{\mathbf{e}\}$ both form (improper) subgroups where $\{\mathbf{e}\}$ is known as the *trivial subgroup*. If given a nonempty subset H of G and asked to determine if H is indeed a subgroup of G, one need not check all three group axioms. It is sufficient for H to satisfy the following *subgroup criterion*.

Proposition 2.0.2 (Subgroup Criterion)**.** *Let* $\emptyset \neq H \subseteq G$ *and* $a, b \in H$*. If* $ab^{-1} \in H$ *then* $H \leqslant G$*.*

Proof. See [32]. □

Definition 2.0.3. Define a group $G[2] = \{a \in G \,:\, a^2 = \mathbf{e}\}$, this abelian group is called the *Boolean group* of G which is also denoted by $B(G)$. The definition of $G[2]$ makes it obvious that in $G[2]$ every element is its own inverse, i.e. $a^{-1} = a$, $\forall a \in G[2]$. Furthermore $|G[2]|$ is called the 2-rank of G.

Boolean groups play an important role in symmetric colorings.

Definition 2.0.4. $H \leqslant G$ is called a *normal subgroup* of G, denoted by $H \lhd G$, if $gHg^{-1} = H$, $\forall g \in G$.

A *simple group* is a group with only two normal subgroups namely the trivial subgroup and itself. The *centre* $Z(G)$ of G is defined as

$$Z(G) = \{x \in G \,:\, xg = gx \;\forall\; g \in G\;\}.$$

$Z(G)$ is known to be always normal in G. Let $z \in G$, the *centralizer* of z in G is defined as

$$C_G(z) = \{g \in G : zg = gz\;\}.$$

For $x, y \in G$, the *conjugate* of x under y is given by $x^y = y^{-1}xy$. The *normalizer* of G is defined as

$$N_G(H) = \{g \in G \,:\, gHg^{-1} = H\}.$$

Clearly if H is normal in G then $N_G(H) = H$.

2.1 Quotient groups

Definition 2.1.1. Let $H \leqslant G$. Then for $g \in G$, the set $gH = \{gh : h \in H\}$ is called a *left coset* of H in G. The definition for a *right coset* Hg is similar.

The number of left cosets and right cosets in G are known to be equal.

Proposition 2.1.1. *Let $H \leqslant G$ then for $x, y \in G$*

(1) $xH = yH \iff y^{-1}x \in H$,

(2) *if $xH \cap yH \neq \emptyset$, then $xH = yH$ and*

(3) $|xH| = |H| \ \forall x \in G$.

Proof. **(1)** "$\Longrightarrow$" Suppose $xH = yH$, then for any $h_1 \in H$, $\exists h_2 \in H$ such that $xh_1 = yh_2 \Longrightarrow y^{-1}x = h_2h_1^{-1} \in H$.
"$\Longleftarrow$" Suppose $y^{-1}x \in H$ and let $y^{-1}x = h_0$. Now if $a \in xH$, then $a = xh$, for $h \in H \Longrightarrow a = y(y^{-1}x)h = yh_0h = yh_1 \in yH$, then $xH \subseteq yH$. Similarly $yH \subseteq xH$ and **(1)** follows.

(2) Suppose $xH \cap yH \neq \emptyset$, then $\exists a \in xH \cap yH$ such that $xh_1 = a = yh_2$, for $h_1, h_2 \in H$. Then $x = yh_2h_1^{-1} \in yH \Longrightarrow xH \subseteq yH$ and $y = xh_1h_2^{-1} \in xH \Longrightarrow yH \subseteq xH$. Hence $xH = yH$.

(3) Let $h_1, h_2 \in H$ where $h_1 \neq h_2 \Longrightarrow xh_1 \neq xh_2$. This means that multiplying the elements of H by x will produce the same number of elements, i.e., $|xH| = |H|$.

□

Definition 2.1.2. If $H \leqslant G$, then the number of left cosets of H is called the *index* of H in G and is denoted by $[G : H]$.

Suppose $H \lhd G$, then the set of all cosets of H in G form a group under the operation of coset multiplication $\cdot$ defined as follows;

$$gH \cdot g'H = gg'H.$$

It is trivial to show that this operation is associative, see [33] for verification. We denote this group by G/H and it is referred to as the *quotient* or *factor group*. The identity element of this group becomes the coset H and the inverse of a coset aH is $a^{-1}H$. The order simply follows from the definition of the index, i.e. $|G/H| = [G : H]$.

We conclude this section by stating and proving what is known as the first major result in group theory.

Theorem 2.1.2 (Lagrange)**.** *Let G be a finite group and $H \leqslant G$. Then $|H|$ divides $|G|$ and $[G : H] = \frac{|G|}{|H|}$.*

Proof. Let $|G| = n$ and let

$$\{x_1H, x_2H, \cdots, x_nH\}$$

be the family of all left cosets of H in G. Then

$$G = x_1H \cup x_2H \cup \cdots \cup x_nH.$$

Since $G = \{x_1, x_2, \cdots, x_n\}$ and $\mathbf{e} \in H$, by Proposition 2.1.1 **(2)**, for any two cosets x_iH and x_jH, there are only two possibilities;

$$x_iH \cap x_jH = \emptyset \quad \text{or} \quad x_iH = x_jH.$$

It also follows from Proposition 2.1.1 **(3)** that all the cosets have $|H|$ elements. Hence $|G| = |H| + |H| + \cdots + |H| \implies |G| = d|H|$, where $d \in \mathbb{Z}^+$, and the result follows. □

The converse of this theorem is false in general, however when there is a prime factor p of $|G|$ then the theorem holds true and this result is known as Cauchy's theorem.

Theorem 2.1.3 (Cauchy)**.** *Let G be a group and suppose that p is a prime factor of $|G|$. Then there exists a subgroup of order p in G, equivalently G contains an element of order p.*

Proof. See [23]. □

2.2 Homomorphisms and Isomorphisms

Informally, homomorphisms are maps from one group to another that preserve group structure.

Definition 2.2.1. Let $(G, *)$ and $(H, \circ)$ be groups. A *homomorphism* is a map $f : G \longrightarrow H$ such that $\forall g, h \in G$,

$$f(g * h) = f(g) \circ f(h).$$

Proposition 2.2.1. *Given a homomorphism $f : G \longrightarrow H$. Then,*

(1) *$f(\mathbf{e}_G) = \mathbf{e}_H$, where $\mathbf{e}_G$ and $\mathbf{e}_H$ refer to the identity elements in G and H, respectively and*

(2) *for $g \in G$, $f(g^{-1}) = f(g)^{-1}$.*

Proof. The proof of **(1)** and **(2)** are quite trivial and can be found in [33]. □

The *image* of homomorphism f is the set $Imf = \{h \in H \ : \ f(g) = h, \ g \in G\}$. The *kernel* of f is the set of all elements in G that are mapped to the identity element of H, formally, $Kerf = \{g \in G \ : \ f(g) = \mathbf{e}_H\}$.

Example 2.2.1. Let $H \lhd G$ and define $\phi : G \longrightarrow G/H$ by $\phi(g) = gH$ for $g \in G$. Then it is easy to verify that ϕ is a homomorphism with kernel H and it is commonly known as the *natural homomorphism*.

Proposition 2.2.2. *If $f : G \longrightarrow H$ is a homomorphism, then $Kerf \lhd G$.*

Proof. Let $K = Kerf$ and $k \in K$. Then $f(k) = \mathbf{e}_G$. Now $f(gkg^{-1}) = f(g)f(k)f(g^{-1}) = f(g)f(k)f(g)^{-1} = \mathbf{e}_H$, for all $g \in G$, therefore $gkg^{-1} \in K \Longrightarrow gKg^{-1} \leq K$ and clearly $K \leq gKg^{-1}$. Hence $K \lhd G$. □

A homomorphism that is also a bijection is called an *isomorphism*. We say that G is isomorphic to H, denoted by $G \cong H$, if there exists an isomorphism $f : G \longrightarrow H$. An *automorphism* is simply an isomorphism from a group to itself. The set of all automorphisms forms a group, called the automorphism group denoted by $Aut(G) = \{\phi : G \longrightarrow G \ : \ \phi \text{ is an automorphism}\}$.

Theorem 2.2.3 (First Isomorphism Theorem)**.** *Let $f : G \longrightarrow H$ be a homomorphism. Then $Kerf \lhd G$ and $G/Kerf \cong Imf$.*

Proof. It was already shown in the previous proposition that $Kerf \lhd G$. So let $K = Kerf$ and define a homomorphism $\phi : G/K \longrightarrow H$ as follows $\phi(gK) = f(g)$.

(1) ϕ is well-defined:

Suppose $gK = hK \Longrightarrow gh^{-1} \in K$. Then we have that $f(gh^{-1}) = f(g)f(h^{-1}) = \mathbf{e}_H \Longrightarrow f(g) = f(h)$ which shows that $\phi(gK) = \phi(hK)$.

(2) ϕ is a homomorphism:

$\phi(gK \cdot hK) = \phi(ghK) = f(gh) = f(g)f(h) = \phi(gK)\phi(hK)$.

(3) ϕ is bijective:

It is clear that $Im\phi = Imf$ therefore it remains to show that ϕ is injective. So let $\phi(gK) = \phi(hK) \Longrightarrow f(g) = f(h) \Longrightarrow f(h^{-1}g) = \mathbf{e}_H \Longrightarrow h^{-1}g \in K \Longrightarrow gK = hK$. Hence ϕ is an isomorphism.

□

It follows from this theorem that a quotient group and a homomorphic image can be thought of as being equivalent.

2.3 Cyclic groups

Definition 2.3.1. G is called a *cyclic group* of order n if $G = \langle x \rangle$ and $x^n = \mathbf{e}$.

Example 2.3.1. $\mathbb{Z}$ forms a group under addition. In fact $(\mathbb{Z}, +)$ is an infinite cyclic group.

Every cyclic group is abelian and we can easily deduce that a subgroup of an abelian group is normal. Suppose we have two cyclic groups G_1 and G_2 of the same order, then it can easily be shown that $G_1 \cong G_2$. We use C_n to denote the general finite cyclic group of order n, although the notation $\mathbb{Z}_n$ is also used interchangeably.

Theorem 2.3.1 (Fundamental Theorem of Cyclic Groups)**.** *Let $G = \langle x \rangle$ be a cyclic group of order n then;*

(1) *every subgroup of G is also cyclic and*

(2) *for each factor k of n, there exists exactly one subgroup of order k, namely $\langle x^{\frac{n}{k}} \rangle$.*

Proof. See [33]. □

Definition 2.3.2. Let $n \in \mathbb{Z}^+$, the group of integers modulo n forms a finite abelian group under addition, denoted by $\mathbb{Z}_n$, where $|\mathbb{Z}_n| = n$. It is important to note that $\mathbb{Z}_n \cong C_n$.

2.4 Presentations of groups

Sometimes it is advantageous to have a more convenient method of defining a group, one such method is called a *group presentation*.

Definition 2.4.1. Let X be a subset of a group G, then $\langle X \rangle$ is the subgroup of all the elements of G that are expressible as the finite product of elements in X and their inverses. If $G = \langle X \rangle$, then we call the elements in X generators of the group G and we say that X generates G.

Example 2.4.1. When $\langle X \rangle$ consists of only a single element $x \in X$, $\langle X \rangle$ is written as $\langle x \rangle$ and $\langle x \rangle$ is a cyclic group, equivalently if an element x generates a group then $\langle x \rangle$ equals the entire group G.

Definition 2.4.2. A relation in a group G is an equation that the generators of G satisfy.

Definition 2.4.3. Let X be a set of generators of a group G and R be a set of relations that the generators satisfy. If G can be written as

$$G = \langle X \, : \, R \rangle,$$

then this is called a presentation of G.

Example 2.4.2. A cyclic group C_n can be expressed as

$$C_n = \langle x \, : \, x^n = \mathbf{e} \rangle,$$

which is a presentation of C_n with only one generator x and one relation $x^n = \mathbf{e}$.

Example 2.4.3. The non-abelian group of order 8 called the *Quaternion group* whose elements are as follows;

$$Q_8 = \{\pm\mathbf{1}, \pm i, \pm j, \pm k\}$$

where $\mathbf{1}$ is the identity element and the relations between elements in Q_8 are;

(1) $i^2 = j^2 = k^2 = -\mathbf{1}$,

(2) $ij = -ji = k$, $jk = -kj = i$, $ki = -ik = j$ and

(3) $-\mathbf{1}$ commutes with every element in Q_8.

Therefore the presentation of Q_8 is

$$Q_8 = \langle a, b : a^4 = \mathbf{1}, a^2 = b^2, b^{-1}ab = a^{-1} \rangle.$$

Example 2.4.4 (*Presentation of the Dihedral group*)**.** The group of all symmetries of a regular polygon of n sides (n-gon) is called the *Dihedral group*, denoted by D_n. Let us consider the symmetry group of the square which is D_4. Suppose that this square is centered at the origin of the xy-axis so that its sides are parallel to the x and y axes. The symmetries of the square (regular 4-gon) are the reflections about the square's diagonals as well as the (anticlockwise) rotations by 0, $\frac{\pi}{2}$, π and $\frac{3\pi}{2}$ radians. A rotation by 0 radians is clearly the identity $\mathbf{e}$. Let a denote the rotation by $\frac{\pi}{2}$ radians, then a^2 and a^3 become the rotations by π and $\frac{3\pi}{2}$ radians, respectively and $a^4 = \mathbf{e}$. Let b denote the reflection about the y-axis, then $b^2 = \mathbf{e}$ and the following occurs;

(1) ab becomes the reflection about the main diagonal (note that action b takes place first),

(2) a^2b is the reflection about the x-axis and

(3) $a^3b = ba$ is the reflection about the other diagonal.

Hence all the symmetries of the square, or all the elements of D_4, can be written with respect to a and b, which implies that a and b are the generators of D_4 and the equations $a^4 = \mathbf{e}$, $b^2 = \mathbf{e}$ and $a^3b = ba$ are the relations of D_4.

Hence the group D_4 is given by the presentation

$$D_4 = \langle a,\, b \,:\, a^4 = \mathbf{e},\, b^2 = \mathbf{e},\, a^3 b = ba \rangle.$$

The presentation of every Dihedral group is analogous to this one. In general the group D_n of order $2n$, has two generators namely a, which is the rotation about $\dfrac{2\pi}{n}$ and b, which is the reflection about the y-axis. These generators yield the relations $a^n = \mathbf{e}$, $b^2 = \mathbf{e}$ and $ba = a^{n-1}b$. Thus the presentation of D_n is

$$D_n = \langle a,\, b \,:\, a^n = \mathbf{e},\, b^2 = \mathbf{e},\, a^{n-1} b = ba \rangle.$$

Whereas D_n describes the rotational and reflection symmetries of regular n-gons, $\mathbb{Z}_n$ depicts the group of only the rotational symmetries of a regular n-gon.

2.5 Group Actions

Group actions validate the relationship between a group and the symmetries of an object. It associates the elements of the group to specific transformations of the object.

Definition 2.5.1. Let G be a group and A be a nonempty set. G is said to *act* on A if there exists a map $f : G \times A \longrightarrow A$, where $(g,\, a) \longrightarrow g \cdot a$ for $g \in G$, $a \in A$ and $g \cdot a \in A$, such that the following two axioms hold;

(1) $\mathbf{e} \cdot a = a$, $\forall a \in A$ and

(2) for all $g,\, h \in G$ and $a \in A$, $(g \cdot h) \cdot a = g \cdot (h \cdot a)$.

Then the map f is called a *left group action* of G on A, a *right group action* is defined in a similar way. Every group G is known to act on itself if the set $A = G$.

Definition 2.5.2. Let $X = \{1, 2, \cdots, n\}$. The *symmetric group*, denoted by S_n, of degree n is the group of all permutations of X. Therefore the order of S_n is $n!$.

Example 2.5.1. S_3 is the symmetric group of all permutations of the set $\{1, 2, 3\}$. It is the *smallest* nonabelian group in existence and $S_3 \cong D_3$.

Definition 2.5.3. The *alternating group*, denoted by A_n, is the group of all even permutations of set $X = \{1, 2, \cdots, n\}$. The order of A_n is $\frac{n!}{2}$.

Definition 2.5.4. Let G be a group acting on a nonempty set S. For $s \in S$, the *orbit* of s under G is the set

$$Orb(s) = \{g \cdot s \, : \, g \in G\}.$$

The orbits of S materialize into a partition since the relation $s \equiv s'$, defined by $s' = g \cdot s$ for some $g \in G$ and s, $s' \in S$, is an equivalence relation whose equivalence classes form the orbits.

Definition 2.5.5. Let G be a group that acts on a nonempty set S. For $s \in S$, the *stabilizer* of s in G is the set

$$Stab(s) = \{g \in G \, : \, g \cdot s = s\}.$$

It is fairly easy to verify that $Stab(s) \leqslant G$. Firstly $\mathbf{e} \in Stab(s)$ by axiom **(1)**

of a group action. Now let g, $h \in Stab(s)$ then;

$$\begin{aligned} s &= \mathbf{e} \cdot s \\ &= (h^{-1}h) \cdot s \\ &= h^{-1}(h \cdot s) \\ &= h^{-1}s \text{ because } h \cdot s = s. \end{aligned}$$

So $h^{-1} \in Stab(s)$, therefore $(gh^{-1}) \cdot s = g \cdot (h^{-1}s) = g \cdot s = s$. Hence $gh^{-1} \in Stab(s)$ and so by the *subgroup criterion* (Proposition 2.0.2), $Stab(s) \leqslant G$. This is a simple yet useful result since it shows that any set which appears to be a stabilizer in a group action is guaranteed to be a subgroup as well.

Theorem 2.5.1 (Orbit-Stabilizer Theorem)**.** *Suppose G is a group that acts on a finite set S and let Orb(s) and Stab(s) be the orbit and stabilizer of $s \in S$ respectively, then the size of the orbit is the index of the stabilizer i.e.*

$$|Orb(s)| = [G : Stab(s)] = \frac{|G|}{|Stab(s)|}.$$

Proof. Let $Stab(s) = H$. Now let us define a map f from $Orb(s)$ to the set of left cosets of H in G i.e. $f : Orb(s) \longrightarrow G/H$. Take $y \in Orb(s)$. Then $y = g \cdot s$, for some $g \in G$ and define $f(y) = gH$.

f is well-defined:

So $g_1 \cdot s = g_2 \cdot s \Longrightarrow (g_2^{-1}g_1) \cdot s = s \Longrightarrow g_2^{-1}g_1 \in H \Longrightarrow g_1H = g_2H$. It is trivially deducible that f is surjective.

f is injective:

Let $y_1 = g_1 \cdot s$ and $y_2 = g_2 \cdot s$ and suppose $f(y_1) = f(y_2)$.

Then $g_1H = g_2H \implies g_2^{-1}g_1 \in H \implies (g_2^{-1}g_1) \cdot s = s \implies g_1 \cdot s = g_2 \cdot s$. Thus $y_1 = y_2$. Hence f is a well-defined bijective map and the result follows. $\square$

The following lemma is also referred to as the *Orbit-Counting Theorem*.

Lemma 2.5.2 (Burnside)**.** *Let G be a finite group that acts on a set S and for each $g \in G$, let S^g denote the set of elements in S that are fixed by g, expressed as $S^g = \{s \in S \,:\, g \cdot s = s\}$. Define S/G to be set of all orbits of G. Then*

$$|S/G| = \frac{1}{|G|} \sum_{g \in G} |S^g|.$$

Proof.

$$\sum_{g \in G} |S^g| = |\{(g,\, s) \in G \times S \,:\, g \cdot s = s\}| = \sum_{s \in S} |Stab(s)|,$$

but by the Orbit-Stabilizer theorem

$$|Orb(s)| = [G : Stab(s)] = \frac{|G|}{|Stab(s)|},$$

therefore

$$\sum_{s \in S} |Stab(s)| = \sum_{s \in S} \frac{|G|}{|Orb(s)|} = |G| \sum_{s \in S} \frac{1}{|Orb(s)|}.$$

Notice that S is the disjoint union of all its orbits in $S/G \implies$ the sum over S can be broken up into seperate sums over each individual orbit as follows,

$$\sum_{s \in S} \frac{1}{|Orb(s)|} = \sum_{A \in S/G} \sum_{s \in A} \frac{1}{|A|} = \sum_{A \in S/G} 1 = |S/G| \implies \sum_{g \in G} |S^g| = |G| \cdot |S/G|$$

and the result follows. $\square$

2.6 Direct and Semidirect products

One of the simplest and most important ways of constructing a new group from already given ones is by considering the ordered pairs of elements of these groups.

Definition 2.6.1. Let $G_1, G_2 \cdots, G_k$ be groups. The *direct product* of these groups, denoted by $G_1 \times G_2 \times \cdots \times G_k$, forms a group with elements of the form $(g_1, g_2, \cdots, g_k)$, where $g_i \in G_i$ for $i = 1, \cdots, k$. And the group operation on the elements of $G = G_1 \times G_2 \times \cdots \times G_k$ is componentwise multiplication defined as follows;

$$(g_1, g_2, \cdots, g_k)(g_1', g_2', \cdots, g_k') = (g_1 g_1', g_2 g_2', \cdots, g_k g_k').$$

The group axioms for G are easily verifiable, G is also refered to as the *Cartesian product* of the groups $G_1, G_2, \cdots, G_k$. Each G_i $(i = 1, \cdots, k)$, is called a *factor* of G and $|G| = |G_1|\,|G_2| \cdots |G_k|$. Finite abelian groups can be completely determined by the direct product of cyclic groups, this profound result is stated below.

Theorem 2.6.1 (Fundamental Theorem of Finite Abelian Groups)**.** *Let G be a finite Abelian group and let $C_1, C_2, \cdots, C_n$ be unique cyclic groups each of prime power order, then G is isomorphic to the direct product of these cyclic groups, in symbols;*

$$G \cong C_1 \times C_2 \times \cdots C_n.$$

Proof. See [33]. □

Example 2.6.1. The Klein four group, symbolised by V_4, is an abelian group of four elements, whose presentation is

$$V_4 = \langle a,\, b \,:\, a^2 = b^2 = \mathbf{e},\, ab = ba \rangle.$$

It is well-known that $V_4 \cong D_2$ and $V_4 \cong C_2 \times C_2$. In fact, every cyclic group C_n, where $n = pq$, is expressible as a direct product of C_p and C_q as long as p and q are relatively prime.

Definition 2.6.2. Let $\phi : H \longrightarrow Aut(N)$ be a homomorphism defined as follows;

$$\phi(h) = \phi_h,\ \forall h \in H.$$

Where the automorphism $\phi_h : N \longrightarrow N$ is defined as

$$\phi_h(n) = hnh^{-1}.$$

The *(outer) semidirect product* w.r.t ϕ is the set

$$H \ltimes_\phi N = \{(h,\, n) \,:\, h \in H,\, n \in N\},$$

with the operation $*$ on the elements defined as follows;

$$(h_1,\, n_1) * (h_2,\, n_2) = (h_1 h_2,\, n_1 \phi_{h_1}(n_2)).$$

A group G is the (inner) semidirect product of N by H if the following conditions are satisfied;

(1) $G = NH$

(2) $N \lhd G$ and

(3) $H \cap N = \mathbf{e}$.

Then $G = H \ltimes N$ and we say that G is a *split extension* of N by H. H is called the *complement* of N in G and N is called the *normal complement* of H in G. The sets $N' = \{(n, \mathbf{e}_H) : n \in N\}$ and $H' = \{(\mathbf{e}_N, h) : h \in H\}$ are subgroups of G (in fact $N' \lhd G$), where $N' \cong N$, $H' \cong H$ and G is the semidirect product of N' by H'.

Definition 2.6.3. Let A be an abelian group, the *generalised Dihedral group* of A, denoted by $Dih(A)$ is the semidirect product of A and C_2 w.r.t ϕ, symbolically

$$Dih(A) \cong C_2 \ltimes_\phi A.$$

In fact C_2 acts on A by means of $\phi : C_2 \longrightarrow Aut(A)$ which is defined as follows, for $\mathbf{e} \neq g \in C_2$,

$$\phi_g(a) = a^{-1}, \forall a \in A.$$

The semidirect product is utilized when deriving other well-known products such as the *holomorph* and the *wreath product*. For example take $H = Aut(N)$, then $\phi : Aut(N) \longrightarrow Aut(N)$. The (semidirect) product that emerges from this identity map ϕ is called the holomorph of N.

Chapter 3

Combinatorics

This chapter provides the unifying framework for the combinatorial theory covered in the sequel. Partially ordered sets will be introduced together with their terminology, basic properties and some important examples before proceeding to the highpoint of this chapter which is the Möbius inversion theorem. The material proffered corresponds to [38]. Before we venture into the theory associated with posets, let us address an essential combinatorial principle.

3.1 Inclusion-Exclusion

The principle of inclusion-exclusion[1] is an important tool in combinatorics that connects the sizes of two finite sets and their union. Simply stated, if A_1 and A_2 are two finite sets, then

$$|A_1 \cup A_2| = |A_1| + |A_2| - |A_1 \cap A_2|.$$

[1] Also known as the sieve principle

The inclusion-exclusion principle can be extended to more than two sets.

Theorem 3.1.1 (Inclusion-Exclusion principle)**.** *For nonempty finite sets* $A_1, A_2, \cdots, A_n$

$$|A_1 \cup A_2 \cup \cdots \cup A_n| = \sum_{i=1}^{n} |A_i| - \sum_{1 \leqslant i < j \leqslant n} |A_i \cap A_j| + \sum_{1 \leqslant i < j < k \leqslant n} |A_i \cap A_j \cap A_k| - \cdots + (-1)^{n-1} |A_1 \cap \cdots \cap A_n| . \tag{3.1}$$

Proof. Suppose that $x \in A_1 \cup A_2 \cup \cdots \cup A_n$ belongs to precisely m of the individual sets $A_1, A_2 \cdots, A_n$. Since x is counted exactly once on the left-hand side of equation (3.1), it suffices to show that x is counted exactly once on the right-hand side of (3.1). The number of sets of the form $A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}$ that x belongs to is $\binom{m}{k}$, since it is the the number of ways of choosing k of the sets which contain x. This implies that the right hand side of (3.1) is counted

$$\binom{m}{1} - \binom{m}{2} + \cdots + (-1)^{m-1} \binom{m}{m}$$

many times. By the *binomial theorem*, for $m \neq 0$

$$\binom{m}{0} - \binom{m}{1} + \cdots + (-1)^{m} \binom{m}{m} = 0$$

hence

$$\binom{m}{1} - \binom{m}{2} + \cdots + (-1)^{m-1} \binom{m}{m} = \binom{m}{0} = 1.$$

Therefore x is counted exactly once on the right-hand side of (3.1). □

3.2 Partitions

A central idea in combinatorics is the partitioning of a set S into nonempty disjoint subsets.

Definition 3.2.1. Let S be a finite set, a *partition* of S is the set of subsets of S given by $\pi = \{A_1, A_2, \cdots, A_k\}$, with the following conditions;

(1) A_i is nonempty for each i,

(2) $A_i \cap A_j = \emptyset$ if $i \neq j$ and

(3) $A_1 \cup A_2 \cup \cdots \cup A_k = S$.

Each A_i is called a cell of π and a partition into k cells is called a k-partition.

Alternatively, a partition may also be defined with respect to an equivalence relation $\sim$ on a set S, which means that a partition on S is the set of equivalence classes of $\sim$ on S.

3.3 Partially Ordered sets

In order theory, a *partial order* of a set is a binary relation which guarantees that not every pair of elements in the set need to be related, compared to a *total order*, which relates every pair of elements to each other. A *partially ordered set* (abbreviated to *poset*) is composed of a set together with a partial order. The formal definition is given below.

Definition 3.3.1. Let S be a nonempty set, a *poset* is an ordered pair $P = (S, \leq)$, where $\leq$ is a relation on S that satisfies the following conditions, for each $a, b, c \in P$;

(1) $a \leq b$ (*reflexive*),

(2) if $a \leq b$ and $b \leq a$, then $a = b$ (*antisymmetric*),

(3) if $a \leq b$ and $b \leq c$, then $a \leq c$ (*transitive*).

Using the symbol $\leq$ can be confusing so we typically write $P = (S, \leq_R)$ to denote the poset with relation $\leq_R$. We shall now outline some of the basic definitions and results that accompany partially ordered sets.

Definition 3.3.2. Let P be a poset.

(1) Two elements $a, b \in P$ are called *comparable* if $a \leq b$ or $b \leq a$, otherwise they are *incomparable*, denoted as $a \| b$.

(2) The *dual* of P is the poset P^* on the same set of P but with the partial order relation reversed i.e. $a \leq b$ in $P \iff b \leq a$ in P^*.

(3) If $Q \subseteq P$ with the partial order on Q being such that $a \leq b$ in $Q \iff a \leq b$ in P. Then we call Q a *subposet* of P.

(4) An *open interval* in P is a subposet of P defined by the set

$$(a, b) = \{c \in P \,:\, a < c < b\}.$$

(5) Similarly, a *closed interval* in P is a subposet of P defined by the set

$$[a, b] = \{c \in P \,:\, a \leq c \leq b\}.$$

(6) P is called *locally finite* if every interval of P is finite.

(7) A subposet I of P is said to be an *ideal* of P if the following statements hold;

(i) $b \in I$ and $a \leq b \Longrightarrow a \in I$ and

(ii) $a, b \in I \Longrightarrow \exists c \in I$ such that $a \leq c$ and $b \leq c$.

(8) The minimal ideal that contains a given element a is called a *principal ideal* and a is called a *principal element* of the ideal. The principal ideal $\langle a \rangle$ for principal element a is described by the set

$$\langle a \rangle = \{c \in P \,:\, c \leq a\}.$$

(9) Let $a, b \in P$ then b is said to *cover* a if $a < b$ and there exists no element $c \in P$ such that $a < c < b$. This *cover relation* is denoted by $a \lessdot b$.

Finite posets can be represented pictorially by means of a *Hasse diagram*[2].

Definition 3.3.3. Let the elements a, b which belong to finite poset P be drawn as vertices in a plane such that whenever $a \lessdot b$, then b must be located above a with a straight line drawn connecting a and b. This action is repeated for all pairs of elements in P for which the relation $\lessdot$ holds. The resulting diagram is called a *Hasse diagram*.

The same partial order can be represented by differently drawn Hasse diagrams therefore Hasse diagrams may be isomorphic to each other. There are

[2]named after German mathematician Helmut Hasse who popularized their use

several conventions taken to ensure that "good" Hasse diagrams are produced from partial orders, see [28].

Definition 3.3.4. Let P and Q be two posets, the *direct product*

$$P \times Q = \{(a, b) : a \in P \text{ and } b \in Q\},$$

forms a poset with the relation $\leq$ between two elements (a, b), $(a_1, b_1) \in P \times Q$ defined as follows;

$$(a, b) \leq (a_1, b_1) \Longleftrightarrow a \leq a_1 \text{ and } b \leq b_1.$$

Example 3.3.1. The set of all partitions of $[n] = \{1, 2 \cdots, n\}$ is denoted by Π_n. Π_n can be transformed into a poset if the partial order on the elements of Π_n is defined as follows; let $\sigma, \pi \in \Pi_n$ then

$$\pi \leq \sigma \text{ if every cell of } \pi \text{ is contained in every cell of } \sigma.$$

The relation $\leq$ on the partitions of $[n]$ is called a *refinement order*. For example take $n = 7$ and let $\pi = \{136, 25, 4, 7\}$ and $\sigma = \{1346, 257\}$, then $\pi \leq \sigma$ and π is said to be a *refinement* of σ. Note that $|\pi|$ is used to denote the number of cells in partition π, therefore $|\pi| = 4$ and $|\sigma| = 2$.

Example 3.3.2. Let G be a finite group. The set of all subgroups of G, P_G, forms a poset $P = (P_G, \leq_R)$, with the relation $\leq_R$ defined as follows;

$$H \leq_R H' \text{ if } H \subseteq H' \text{ for } H, H' \in P_G.$$

The set of all subgroups of G also forms a *lattice*, denoted by $L(G)$. For more information on this subgroup lattice $L(G)$, consult [39].

Example 3.3.3 (Constructing the Poset of Optimal Partitions (see [24]))**.** Let G be a finite group with π being a partition of G. The *stabilizer* and the *center* of π are defined respectively as

$$St(\pi) = \{g \in G : \forall x \in G,\ x \text{ and } xg^{-1} \text{ belong to the same cell of } \pi\} \text{ and}$$
$$Z(\pi) = \{g \in G : \forall x \in G,\ x \text{ and } gx^{-1}g \text{ belong to the same cell of } \pi\}.$$

A partition π of G is called an *optimal partition* if $\mathbf{e} \in Z(\pi)$ and for all partitions σ of G, where $St(\sigma) = St(\pi)$ and $Z(\sigma) = Z(\pi)$, $\pi \leq \sigma$ (where $\leq$ is the refinement order). These optimal partitions of G form a poset under the relation $\leq$, which is understood to be the poset of pairs of subsets of G, $(St(\pi), Z(\pi))$, for π a partition of G and $\mathbf{e} \in Z(\pi)$. Creating the poset of optimal partitions of G begins with the *finest partition* $\pi = \{a,\ a^{-1} : a \in G\}$ thereafter we use π to calculate the remaining optimal partitions using the following proposition.

Proposition 3.3.1. *Let π be an optimal partition of G and suppose $X \subseteq G$. Let π_1 be the finest partition of G such that $\pi \leq \pi_1$ and $X \subseteq St(\pi_1)$, also let π_2 be another finest partition of G such that $\pi \leq \pi_2$ and $X \subseteq Z(\pi_2)$. Then π_1 and π_2 are also optimal partitions.*

Proof. This is trivially deducible from the definition of an optimal partition. □

3.4 The Möbius Function

Let Z_1 be the set of all integers greater than or equal to 1 and suppose that we have in our possession two functions $f : Z_1 \longrightarrow \mathbb{Z}$ and $g : Z_1 \longrightarrow \mathbb{Z}$ that satisfy,

$$g(n) = \sum_{d|n} f(d)$$

where $d|n$ means that d divides n. It is possible to derive a formula for $f(n)$, from the one stated above, through a process called *Möbius inversion*, named after its creator August Ferdinand Möbius[3] who proposed it in 1832. It entails finding a function μ such that the following theorem is satisfied.

Theorem 3.4.1. *For $n \in Z_1$, there exists a function $\mu : Z_1 \longrightarrow \mathbb{Z}$ such that*

$$g(n) = \sum_{d|n} f(d) \iff f(n) = \sum_{d|n} \mu(\tfrac{n}{d}) g(d).$$

Proof. Refer to [12]. □

The function μ is found to be

$$\mu(n) = \begin{cases} 1, & \text{for } n = 1 \\ 0, & \text{if there exists prime } p \text{ such that } p^2|n \\ (-1)^k, & \text{if } n = p_1 p_2 \cdots p_k \text{ where } p_1,\, p_2 \cdots,\, p_k \text{ are distinct primes.} \end{cases}$$

The classical number theoretic Möbius inversion formula, stated above, has an assortment of variations. We will now examine one that is specific to poset theory.

[3]Gauss was actually the first to conceive of it 30 years prior to Möbius.

Definition 3.4.1. Let P be a locally finite poset and consider the set

$$I(P) = \{\alpha : P \times P \longrightarrow \mathbb{R} \,:\, \alpha(x,\, y) = 0 \text{ if } x \not\leq y\}.$$

For $\alpha,\, \beta \in I(P)$ the two operations on $I(P)$ are defined as follows.

(1) Addition

$$(\alpha + \beta)(x,\, y) = \alpha(x,\, y) + \beta(x,\, y).$$

(2) Multiplication (also called *convolution* which is associative)

$$(\alpha * \beta)(x,\, y) = \sum_{x \leq z \leq y} \alpha(x,\, z)\beta(z,\, y). \tag{3.2}$$

$I(P)$ forms the *incidence algebra* over $\mathbb{R}$ and if $\alpha \in I(P)$, then α is referred to as an *incidence function.*

The multiplicative identity of $I(P)$ is,

$$\delta(x,\, y) = \begin{cases} 1, & \text{if } x = y \\ 0, & \text{if } x \neq y \end{cases} \tag{3.3}$$

which is identifiable as the *Kronecker delta.*

Proposition 3.4.2. *The multiplicative inverse of a function $\alpha \in I(P)$ exists $\iff \alpha(x,\, x) \neq 0$ for $x \in P$.*

Proof. "$\Longleftarrow$" Suppose $\alpha(x,\, x) \neq 0$, then we can define $\alpha^{-1}(x,\, x) = \dfrac{1}{\alpha(x,\, x)}$.

For $x < y$, from (3.2), we get

$$\alpha^{-1}(x,\, y) = -\frac{1}{\alpha(x,\, x)} \sum_{x<z\leq y} \alpha(x,\, z)\alpha^{-1}(z,\, y)$$

and then by simple computations it is fairly easy to show that

$$\alpha^{-1} * \alpha = \delta = \alpha * \alpha^{-1}.$$

"$\Longrightarrow$" Suppose α is invertible. Then α^{-1} exists and for all x

$$(\alpha * \alpha^{-1})(x,\, x) = \alpha(x,\, x)\alpha^{-1}(x,\, x) = 1$$

by the definition of $*$, hence $\alpha(x,\, x) \neq 0$. □

Definition 3.4.2. The function in $I(P)$

$$\zeta(x,\, y) = \begin{cases} 1, & \text{if } x \leq y \\ 0, & \text{if } x \nleq y \end{cases} \tag{3.4}$$

is called the *zeta function*.

By the above proposition, the multiplicative inverse of the zeta function exists in $I(P)$ and is known as the *Möbius function* μ of P i.e., $\zeta^{-1} = \mu$. We define μ recursively by the conditions below

$$\mu(x,\, y) = \begin{cases} 1, & \text{if } x = y \\ 0, & \text{if } x \nleq y \\ -\sum\limits_{x\leq z<y} \mu(x, z), & \text{if } x < y. \end{cases} \tag{3.5}$$

Theorem 3.4.3 (The Möbius Inversion Formula)**.** *Let P be locally finite poset with Möbius function μ and let $\alpha : P \longrightarrow \mathbb{R}$, $\beta : P \longrightarrow \mathbb{R}$. Assume that all the principal ideals of P are finite, then*

$$\beta(x) = \sum_{y \leq x} \alpha(y) \Longleftrightarrow \alpha(x) = \sum_{y \leq x} \mu(y, x)\beta(y).$$

Proof. "$\Longrightarrow$"

$$\begin{aligned}
\sum_{y \leq x} \mu(y, x)\beta(y) &= \sum_{y \leq x} \mu(y, x) \left(\sum_{z \leq y} \alpha(z) \right) \\
&= \sum_{z \leq y \leq x} \mu(y, x)\alpha(z) \\
&= \sum_{z \leq x} \alpha(z) \left(\sum_{z \leq y \leq x} \mu(y, x) \right) \\
&= \alpha(x)
\end{aligned}$$

It is possible to show that "$\Longleftarrow$" is true by using exactly the same computational argument. □

Chapter 4

Ramsey Theory

"*complete disorder is impossible*"

This philosophical quote, attributed to Theodore Motzkin, thoroughly expresses the governing principle behind Ramsey theory. Although Ramsey theory is thought to have emerged with Frank Plumpton Ramsey's[1] 1928 paper "*On a problem of formal logic*" [31], which proved his eponymous theorem, several Ramseyan-type results predate his work. The first instance of Ramseyan mathematics is widely known to be *Hilbert's cube lemma* which appeared in 1892 in the paper [22].

The second is attributed to Issai Schur, who in his 1916 paper [34], showed that in any finite coloring of $\mathbb{N}$, a monochromatic solution x, y and z to the equation $x + y = z$, must exist. The generalisation of this result was published in 1933 by Schur's PhD student Richard Rado in his thesis [29].

[1] Ramsey was also accomplished in the fields of economics and philosophy before his untimely death at age 26.

The basic idea of Rado's theorem involves writing $x + y - z = 0$, so that this equation becomes a homogeneous linear equation. Rado's theorem produces conditions that a system of such equations must satisfy in order for a monochromatic solution in any r-coloring of $\mathbb{N}$, for $r \in \mathbb{N}$, to exist.

In his 1927 paper [41], Bartel Leendert van der Waerden proved that given any finite coloring of $\mathbb{N}$ there exists a k-term monochromatic arithmetic progression. Van der Waerden's theorem became one of the pioneering results of modern Ramseyan mathematics. Under the tutelage of famous mathematicians such as Paul Erdös, who is arguably the most prolific mathematician of the twentieth century, Ramseyan mathematics has only recently emerged into the fully realized mathematical field named Ramsey theory.

This chapter explores Ramsey theory in an explicit manner by first providing a preliminary section, which serves to introduce notation and other foundational requirements of Ramsey theory, before proceeding to state and prove the classical theorems discussed above. These statements and their proofs are freely adapted from a number of rudimentary texts such as [16], [17] and [25].

4.1 Preliminaries

The following theorem is a generalised version of the *basic pigeonhole principle*, a much celebrated result in mathematics. The infinite version states that if we partition an infinite set into a finite number of sets, then at least one of these sets must contain an infinite number of elements. Given below is a more formalised assertion of this principle.

Theorem 4.1.1 (Generalised Pigeonhole Principle)**.** *Suppose we have k number of elements and r number of sets. If the number of elements partitioned into r sets is greater than rk, then there exists a set that contains a number of elements greater than k.*

Proof. We take a set S such that the cardinality of the set S is greater than rk, $|S| > rk$. Partition S into r subsets so that $S = S_1 \cup S_2 \cup \cdots \cup S_r$. It is required to show that there exists a set S_i such that $|S_i| > k$ for some i. We assume that $|S_i| \leq k$, for all $i = 1, \cdots r$. Now

$$|S| = |S_1 \cup \cdots \cup S_r| = |S_1| + \cdots + |S_r| = \sum_{i=1}^{r} |S_i| \leq rk$$

This is clearly a contradiction therefore there must exist at least one i for which $|S_i| > k$. □

Definition 4.1.1. Let S be a set. The function $\chi : S \longrightarrow [r]$ is called an *r-coloring* of S.

Definition 4.1.2. Suppose a coloring χ is constant on a set S, then χ is said to be *monochromatic* on S.

Example 4.1.1. Take a set $S = [1, 7]$ and let $\chi : [1, 7] \longrightarrow \{1, 2\}$ be defined by $\chi(1) = \chi(3) = \chi(5) = \chi(7) = 1$ and $\chi(2) = \chi(4) = \chi(6) = 2$. Therefore χ is a 2-coloring of S that is monochromatic on $\{1, 3, 5, 7\}$ in color 1 and on $\{2, 4, 6\}$ in color 2.

The example above illustrates that an r-coloring can be thought of as a partition of a set S into r subsets, namely $S_{\chi(x)}$ where $x \in S$.

Definition 4.1.3. An *arithmetic progression* (abbreviated to AP) is a sequence of integers in which the difference, $q \in \mathbb{Z}^+$, between each pair of consecutive terms is constant. A k-term such sequence is of the form x, $x + q$, $x + 2q$, $\cdots$, $x + (k - 1)q$, for x, $k \in \mathbb{Z}$.

Definition 4.1.4. A k-set is a set containing k elements and for a set S, $[S]^k$ denotes the set of all k-sets contained in S i.e.

$$[S]^k = \{S_1 \subseteq S \, : \, |S_1| = k\}.$$

Our next objective is to supply a host of definitions pertaining to graph theory since Ramsey's theorem is commonly stated in terms of *graphs*.

Definition 4.1.5. A *graph* is an ordered pair $G = (V, E)$ consisting of a set of points V called *vertices* and a set E of *edges*, which are pairs of distinct vertices that are connected by straight lines.

Two vertices of a graph $G = (V, E)$ are called *adjacent* if they are connected by an edge whereas two edges of G are called *incident* if they share a vertex.

Definition 4.1.6. The *order* of a graph $G = (V, E)$ is the number of vertices i.e. the cardinality of V. The order of G is denoted by n.

Definition 4.1.7. Given a graph $G = (V, E)$. A *subgraph* of G is an ordered pair $G' = (V', E')$, where $V' \subseteq V$ and $E' \subseteq E$.

Definition 4.1.8. A graph of n vertices (or of order n) is called a *complete* graph if each pair of distinct vertices is connected by an edge, a complete graph of n vertices is denoted by K_n.

Definition 4.1.9. Suppose each edge of a graph G is assigned a color such that adjacent edges are differently coloured. Then G is called a *(properly) edge-colored*, or just *colored* graph. If graph G yields a subgraph all of whose edges is the same color, then it is referred to as a *monochromatic subgraph*.

Definition 4.1.10. A graph whose edges have been colored with r different colors is called an *r-colored graph*.

Definition 4.1.11. If graph G is colored monochromatically in color c then G is called *c-monochrome*.

Definition 4.1.12. Let the complete graph with a *countably infinite* set of vertices V be denoted by $K_{\mathbb{N}}$. A countably infinite set simply means a set in which a one-to-one mapping exists between the elements in the set and $\mathbb{N}$. We will take $V = \mathbb{N}$. Then $K_{\mathbb{N}}$ is called a countably infinite (or just infinite) complete graph.

Definition 4.1.13. A *hypergraph* is an ordered pair $H = (V, E)$ where V is a set of *vertices* and E is a set of nonempty subsets of V called *hyperedges*.

Hypergraphs are considered to be a generalization of the standard graph since hypergraphs are graphs with edges that connect an arbitrary number of vertices whereas graphs connect pairs of vertices.

Definition 4.1.14. Given a hypergraph $H = (V, E)$, a *subhypergraph* of H is an ordered pair $H' = (V', E')$, where $V' \subseteq V$ and $E' \subseteq E$.

Definition 4.1.15. The *chromatic number* of a hypergraph $H = (V, E)$, denoted by $\chi(H)$, is the least integer t such that there exists a function $\gamma : V \longrightarrow [t]$ for which there exists no $e \in E$ and $i \in [t]$ such that $e \subseteq \gamma^{-1}(i)$.

The above definition is somewhat opaque therefore we visualise the function γ to be a t-coloring of the vertices of H. We take an edge $e \in E$ and suppose all the vertices of e have the same color, then e is called monochromatic. Then $\chi(H) = t$ translates to being the least integer t for which there exists a t-coloring of V such that there exists no monochromatic edge in E.

We conclude this section by proving the following combinatorial *compactness principle* of which there are many variants, here it is asserted in terms of hypergraphs, which is its most common form. The proof is considered only when V is countable for hypergraph $H = (V, E)$, proof for arbitrary V makes use of *Tychonoff's theorem* or a suitable equivalent such as the *Axiom of choice* and is found in [17].

Theorem 4.1.2 (Compactness Principle Version 1)**.** *Let $H = (V, E)$ be a hypergraph such that E is finite and $\chi(H) > t$. Then there exists a finite subhypergraph H' of H such that $\chi(H') > t$.*

Proof. Assume that V is countable and w.l.o.g assume that $V = \mathbb{N}$. Define a subhypergraph of H to be $H_n = (V_n, E_n)$, for $n \in \mathbb{N}$ where

$$V_n = [n] \text{ and } E_n = \{e \in E : e \subseteq [n]\}.$$

Let $\chi(H_n) \leq t \, \forall n \in \mathbb{N}$. Then (by the definition of $\chi(H_n)$) there exists a function $\gamma_n : [n] \longrightarrow [t]$ such that no monochromatic edge exists in H_n. Now consider the images $\gamma_n(1)$, where $n \in \mathbb{Z}^+$. We may deduce by the pigeonhole principle that there exists some $i_1 \in [t]$ that occurs an infinite number of times in H. We argue using proof by contradiction as follows.

(1) Define a map $\gamma^*(1) = i_1$ and let $n_1 < n_2 < \cdots$ be the indices such that $\gamma^*_{n_i}(1) = i_1$.

Now consider the images $\gamma_{n_i}(2)$, where $i \geq 2$. As in the previous case, there exists some $i_2 \in [t]$ that occurs an infinite number of times in H (by the pigeonhole principle).

(2) Define a map $\gamma^*(2) = i_2$ and let $n'_1 < n'_2 < \cdots$ be the subsequence of the n_i such that $\gamma^*_{n'_i}(2) = i_2$.

Lastly consider the images $\gamma_{n'_i}(3)$, where $i \geq 3$. Again there exists some $i_3 \in [t]$ that occurs an infinite number of times in H (by the pigeonhole principle).

(3) Define a map $\gamma^*(3) = i_3$ and let $n''_1 < n''_2 < \cdots$ be the subsequence of the n'_i such that $\gamma^*_{n''_i}(3) = i_3$.

By applying this argument[2] (or steps) indefinitely, we may define a map $\gamma^* : V \longrightarrow [t]$ such that no monochromatic edge exists in H i.e. $\forall n$, $\exists n'$ such that $\gamma^*(i) = \gamma_{n'}(i)$. This statement implies, by the definition of $\chi(H)$, that $\forall e \in E$ and $i \in [t]$, $e \nsubseteq \gamma^{*^{-1}}(i)$. This is a contradiction to $\chi(H) > t$ and hence the theorem follows. □

[2]This argument is an application of the *König infinity lemma*.

The same principle is restated in terms of sets below.

Theorem 4.1.3 (Compactness Principle Version 2)**.** *Let r, $k \in \mathbb{N}$ and let $\mathcal{S}$ denote the family of all finite subsets of $\mathbb{N}$. Suppose that for any r-coloring of $[\mathbb{N}]^k$ there exists an $S \in \mathcal{S}$ such that $[S]^k$ is monochromatic. Then there exists an n_0 such that for any r-coloring of $[n]^k$, $n \geq n_0$, there exists an $S \in \mathcal{S}$, $S \subseteq [n]$ such that $[S]^k$ is monochromatic.*

Proof. See [17]. □

4.2 Ramsey's Theorem

This section is an exposition of the various forms of Ramsey's theorem which is best interpreted in a graph coloring setting, allowing for the complexity of the theorems to be easily pictured. Ramsey's theorem possesses many different constructions, we first consider a graph theoretical construction of which we present three cases; the two color case, its generalisation to a finite number of colors and lastly the infinite graph case. After that, we progress to the finite and infinite versions of Ramsey's theorem for arbitrary sets.

Theorem 4.2.1 (Ramsey's Theorem for Two Colors)**.** *For a, $b \geq 2$, there exists a least positive integer $R(a, b) = n$ such that every 2-coloring of K_n, with the colors green and red, yields either a green-monochrome K_a or a red-monochrome K_b subgraph.*

Proof. It is trivially deducible that $R(a, 2) = a$ and $R(2, b) = b$, $\forall a, b \geq 2$ (please consult [6] if unclear). We will prove this statement using induction on the sum $a + b$. For $a + b < 5$, it is trivial. Let $a + b = 5$ then $R(2, 3) =$

$R(3,\,2) = 3$ and we are done. So suppose $a + b \geq 6$, for $a,\, b \geq 3$ and assume $R(a-1,\, b)$ and $R(a,\, b-1)$ exist.

Claim 1:

$$R(a,\, b) \leq R(a-1,\, b) + R(a,\, b-1).$$

Let $R(a-1,\, b) + R(a,\, b-1) = n$ and consider a 2-coloring of K_n (in green and red). Choose a vertex of K_n, say x. We know that there are $n-1$ edges from x to the other vertices. Let A be the cardinality of the set of all green edges from x and B be the cardinality of the set of all red edges from x. Then $A + B = n - 1$.

Claim 2:

$$\text{Either } A \geq R(a-1,\, b) \text{ or } B \geq R(a,\, b-1).$$

Suppose $A < R(a-1,\, b)$ and $B < R(a,\, b-1)$, then $A + B \leq (R(a-1,\, b) - 1) + (R(a,\, b-1) - 1) = n - 2$, which is a contradiction therefore claim 2 is indeed true. We may now assume, w.l.o.g, that $A \geq R(a-1,\, b)$. Let V_x be the set of vertices connected to x by a green edge. Clearly $|V_x| = A \geq R(a-1,\, b)$, by definition. By our inductive hypothesis K_{V_x} yields either a green-monochrome K_{a-1} subgraph or a red-monochrome K_b subgraph. If the latter is true then we are done, so suppose the former statement is true. Then if we connect vertex x to each vertex of K_{a-1} we obtain a green-monochrome K_a subgraph. In both cases we obtain either a green-monochrome K_a subgraph or a red-monochrome K_b subgraph and hence by induction the proof is complete.

□

$R(a, b)$ is called a *Ramsey number*, the formal definition is given below.

Definition 4.2.1. A Ramsey number $R(a, b) = n$, is the order of the smallest complete graph K_n which, when 2-colored with colors c_1 and c_2, yields either a c_1-monochrome K_a or a c_2-monochrome K_b subgraph.

Ramsey numbers are intrinsically difficult to find. We will revisit Ramsey numbers in the next section but before we do so, here is a generalisation of the above theorem to a finite number of colors.

Theorem 4.2.2 (Finite Version of Ramsey's Theorem for Graphs)**.** *For $a_1, a_2, \cdots, a_r \in \mathbb{Z}^+$, there exists a least positive integer $R(a_1, a_2, \cdots, a_r) = n$ such that every r-coloring of K_n yields an i-monochrome K_{a_i} for color i, where $i = 1, 2, \cdots, r$.*

Proof. As in the previous case, we shall proceed by induction but this time on the number of colors r. For $r = 2$ this theorem becomes Theorem 4.2.1. Assume that this statement is true for $r - 1$, where $r > 2$ and make the following claim.

Claim:

$$R(a_1, a_2, \cdots, a_r) \leq R(a_1, \cdots, a_{r-2}, R(a_{r-1}, a_r)).$$

Let $R(a_1, \cdots, a_{r-2}, R(a_{r-1}, a_r)) = m$, it exists by the inductive hypothesis. Consider an r-colored graph G of order m. We use the following "*color blindness*" argument in which we assume that color r is the same as color $r - 1$, therefore G is $(r - 1)$-colored. By the inductive hypothesis, G yields either a j-monochrome K_{a_j} subgraph for some color j, $1 \leq j < r - 1$, or a $K_{R(a_{r-1}, a_r)}$ subgraph, monochromatic in color r which is the same as color

$r-1$. If the former statement is true then we are done and therefore consider the latter case. Now from the definition of the Ramsey number $R(a_{r-1}, a_r)$, either a $(r-1)$-monochrome $K_{a_{r-1}}$ or a r-monochrome K_{a_r} exists and hence the result follows. □

Definition 4.2.2. A *generalized Ramsey number* $R(a_1, a_2, \cdots, a_r) = n \in \mathbb{Z}^+$ is the smallest n such that no matter how any r-coloring of K_n is defined, there is some color i such that there exists an i-monochrome subgraph K_{a_i}.

The theorem above proves the existence of such numbers. We will conclude our exposition of graph Ramsey theory by proving Ramsey's theorem for a finite coloring of an infinite graph.

Theorem 4.2.3 (Infinite Version of Ramsey's Theorem for Graphs)**.** *Let $r \in \mathbb{Z}^+$, then every r-coloring of $K_{\mathbb{N}}$ yields an countably infinite complete monochromatic subgraph.*

Proof. Let χ be an r-coloring of $K_{\mathbb{N}}$. Define the set of all vertices of $K_{\mathbb{N}}$ to be $V_1 = \mathbb{N}$ and choose a vertex $v_1 \in V_1$. Regard all the vertices incident to v_1. Since there are a finite number of colors r, an infinite number of these edges have the same color (by the pigeonhole principle), say $c_1 \in [r]$. Now define

$$V_2 = \{v \in V_1 : v_1 < v, \chi(\{v_1, v\}) = c_1\}.$$

where $\{v_1, v\}$ denotes the edge connecting vertices v_1 and v. Take v_2 to be the first vertex in V_2 and regard all the edges incident to v_2. An infinite number of these edges must have the same color, say $c_2 \in [r]$. Define

$$V_3 = \{v \in V_2 : v_2 < v, \chi(\{v_2, v\}) = c_2\}.$$

and take v_3 to be the first vertex in V_3. By continuing the above process indefinitely, we get three infinite sequences, namely

(1) $V_1, V_2, V_3, \cdots,$

(2) $v_1, v_2, v_3, \cdots$ and

(3) $c_1, c_2, c_3, \cdots$

and the following events occur

(1) $V_1 \supsetneqq V_2 \supsetneqq V_3 \cdots,$

(2) v_i is the first vertex in V_i and

(3) for all $v \in V_k$, $k > i$, the edge $\chi(\{v_i, v\}) = c_i$, for $c_i \in [r]$.

An infinite number of colors from the sequence $c_1, c_2, c_3, \cdots$ must be the same, say $c_i = c \in [r]$. Define the set $K_c = \{v_i : c_i = c\}$, then K_c forms an infinite monochromatic complete subgraph of $K_{\mathbb{N}}$. □

We shall now embark on a set theoretic interpretation of Ramsey's theorem.

Theorem 4.2.4 (Infinite Version of Ramsey's Theorem for Sets)**.** *For all $k, r \in \mathbb{N}$ and for every r-coloring $\chi : [\mathbb{N}]^k \longrightarrow [r]$ of the k-element subsets of $\mathbb{N}$, there exists an infinite subset $S \subseteq \mathbb{N}$ such that all its k-element subsets have the same colors.*

Proof. We break the proof into several cases.

Case 1:

For $k = 1$ this theorem becomes an infinite version of the basic pigeonhole principle in which the 1-subsets are simply the elements of $\mathbb{N}$, which is an

infinite set. Therefore any r-coloring of $\mathbb{N}$ will produce a color, say c_1, that occurs an infinite number of times. The c_1-colored elements are the elements contained in the infinite set S and we are done.

Case 2:

Let $k = 2$. We draw the elements of the set $[\mathbb{N}]^2$ as edges of the complete graph of $\mathbb{N}$ vertices, $K_{\mathbb{N}}$, and let $\chi : [\mathbb{N}]^2 \longrightarrow \{1, \cdots, r\}$ be any r-coloring of $K_{\mathbb{N}}$. Call the set of points $Y_0 = \mathbb{N}$ and then fix a point, say $y_0 \in Y_0$. We may deduce by the basic pigeonhole principle that infinitely many of all the edges that connect y_0 with the set of points $Y_0 - \{y_0\}$ are colored with the same color, say c_0. Now call that set of points

$$Y_1 = \{y \in Y_0 - \{y_0\} \, : \, \chi(y_0, y) = c_0\}.$$

Fix a point in this set $y_1 \in Y_1$ and repeat the above process. By the basic pigeonhole principle, infinitely many of all the edges that connect y_1 with the set of points $Y_1 - \{y_1\}$ are colored with the same color, say c_1 and call that set of points

$$Y_2 = \{y \in Y_1 - \{y_1\} \, : \, \chi(y_1, y) = c_1\}.$$

By repeating this process indefinitely, we obtain an infinite set of points $E = \{y_0, y_1, \cdots\}$. Now the color of any pair of edges $\{e, e'\} \in [E]^2$ that connects the points of E is wholely dependent on $\min\{e, e'\}$. Hence we may define a new coloring

$$\chi^*(e) = \chi(\{e, e'\}) \text{ where } e' > e \in E.$$

By the pigeonhole principle, there exists an infinite subset $S \subseteq E$ that is monochromatic w.r.t χ^* i.e. $\chi^*(s)$ is the same, for all $s \in S$. From the definition of χ^*, this simply means that all $\{s, s'\} \in [E]^2$ have the same color under χ and hence Ramsey's theorem for $k = 2$ is proved.

Case 3:

Let $k = 3$, we proceed by using the same technique above. Call the set of points $Y_0 = \mathbb{N}$ and then fix a point, say $y_0 \in Y_0$. Using the r-coloring $\chi : [\mathbb{N}]^3 \longrightarrow \{1, \cdots, r\}$ we define another r-coloring χ_0 on the pairs of $Y = Y_0 - \{y_0\}$ by $\chi_0(\{y, y'\}) = \chi(\{y_0, y, y'\})$, where $y, y' \in Y$. By case 2, Y contains an infinite subset, say Y_1, monochromatic in χ_0 i.e. $\chi_0(y, y') = c_1$, for $y, y' \in Y_1 (y \neq y')$. Fix a point $y_1 \in Y_1$ such that $y_1 > y_0$. Any r-coloring χ_1 on the pairs of $Y' = Y_1 - \{y_1\}$ induces another r-coloring $\chi_1(\{y, y'\}) = \chi(\{y_1, y, y'\})$, where $y, y' \in Y'$. Again by case 2, Y' contains an infinite subset, say Y_2, monochromatic in χ_1 i.e. $\chi_1(y, y') = c_2$, for $y, y' \in Y_2 (y \neq y')$.

By repeating this process indefinitely we obtain an infinite set of points $E = \{y_0, y_1, \cdots\}$ such that the color of any set of edges $\{e, e', e''\} \in [E]^3$ that connects the points of E, is wholely dependent on $\min\{e, e', e''\}$. Hence we may define a new coloring

$$\chi^*(e) = \chi(\{e, e', e''\}) \text{ where } e'' > e' > e \in E.$$

By the pigeonhole principle there must exist an infinite subset $S \subseteq E$ that is monochromatic w.r.t χ^* i.e. $\chi^*(s)$ is the same, $\forall s \in S$. From the definition

of χ^*, this simply means that all $\{s, s', s''\} \in [E]^3$ have the same color under χ and hence Ramsey's theorem for $k = 3$ is proved. The case for any general value of k can be shown by using the same argument and hence the theorem follows. □

Theorem 4.2.5 (Finite Version of Ramsey's Theorem for Sets)**.** *For all $k, l, r \in \mathbb{N}$ there exists a $R(k, l, r) \in \mathbb{N}$ called a Ramsey number such that for all $n \geq R(k, l, r)$ and any r-coloring of the k-subsets of $[n]$,*
$\chi : [n]^k \longrightarrow \{1, \cdots, r\}$, there exists an l-subset of $[n]$ with all of its k-subsets being monochromatic.

Proof. This statement follows immediately from the infinite version of Ramsey's theorem (Theorem 4.2.4) coupled with the second version of the compactness principle (Theorem 4.1.3). □

4.3 Ramsey Numbers

It was mentioned previously that calculating Ramsey numbers is a gruelling task therefore it is easier to find bounds for them instead of exact values. This section provides a requisite table of known Ramsey numbers. This amusing quote, paraphrased from legendary Hungarian mathematician Paul Erdös, who is noted for his major contributions to combinatorics and graph theory, conveys our ineptitude at finding Ramsey numbers.

> "*Suppose aliens invade the earth and threaten to obliterate it in a year's time unless human beings can find the Ramsey number $R(5, 5)$. We could marshal the world's best minds and fastest*

computers, and within a year we could probably calculate the value. If the aliens demanded the Ramsey number $R(6, 6)$, however, we would have no choice but to launch a preemptive attack."

A summary of known values and bounds for Ramsey numbers is showcased in Table 4.1. This table is extracted from Radziszowski's survey of Ramsey numbers [30], which remains the most up-to-date list of Ramsey values that many mathematicians have laboured to find and which also presents bounds for generalized Ramsey numbers whereas we only consider two color Ramsey numbers. Please observe that in the following table $x - y$ is used to denote the interval $[x, y]$, for positive integers x and y.

a \ b	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**
3	6	9	14	18	23	28	36	40 – 43	46 – 51	52 – 59	59 – 69	66 – 78	73 – 88
4	-	18	25	35– 41	49– 61	56– 84	73– 115	92 – 149	97 – 191	128– 238	133– 291	141– 349	153 – 417
5	-	-	43– 49	58– 87	80– 143	101– 216	125– 316	143– 442	157– 1000	181– 1364	205– 1819	233– 2349	261 – 3059
6	-	-	-	102– 165	113– 298	127– 495	169– 780	179– 1171	253– 3002	262– 4367	317– 6187	317– 8567	401 – 11627
7	-	-	-	-	205– 540	216– 1031	233– 1713	232– 2826	405– 8007	416– 12375	511– 18563	511– 27131	511 – 38759
8	-	-	-	-	-	282– 1870	317– 3583	377– 6090	377– 19447	377– 31823	817– 50387	817– 77519	861 – 116279
9	-	-	-	-	-	-	565– 6588	580– 12677	22325	38832	64864	-	-
10	-	-	-	-	-	-	-	798– 23556	45881	81123	-	-	1265

Table 4.1: Known values and bounds for Ramsey numbers $R(a, b)$.

4.4 Van der Waerden's Theorem

This section is devoted to van der Waerden's definitive theorem. Although first published in 1927, there is evidence that this result was first conjectured

by Issai Schur, a few years prior to this, whilst he was working on the distribution of quadratic residues modulo p thereafter van der Waerden became aware of Schur's result through a fellow student Pierre Baudet and referred to it as Baudet's conjecture. Broadly speaking, it suggests that for any coloring of the positive integers $\mathbb{Z}^+$, an arbitrarily long AP is guaranteed to be found. Various different proofs of this result can be found, however in this section we present a purely combinatorial one which complies with van der Waerden's original proof in [41], using the "color-focusing technique" adapted from [42] and [40].

Theorem 4.4.1 (Van der Waerden)**.** *Suppose $k, r \in \mathbb{Z}^+$, then there exists a least positive integer $W(k, r) = n$, such that any r-coloring of $[n]$ yields a monochromatic k-term AP.*

$W(k, r) = n$ is called a *classical van der Waerden number* and naturally this theorem proves the existence of such values. The proof of the above theorem proceeds via double induction on k and r and is postponed until after we address the following proposition, which exhibits this inductive argument.

Proposition 4.4.2. $W(3, 2) \leq 325$.

Proof. Let $S = [1, 325]$. We first 2-color S and then break up S into individual sets[3] each of order 5 obtaining 65 sets in total i.e. $S = [1, 5] \cup [6, 10] \cup \cdots \cup [321, 325]$. Name these sets $S_1, S_2, \cdots, S_{65}$, respectively. Since each of the S_i's contains 5 elements, the elements can be 2-colored in one of $2^5 = 32$ ways. Therefore by the pigeonhole principle, at least two of the first 33 sets

[3]Also referred to as blocks.

are colored in the same way. Name these sets S_a and S_{a+d}. Since every element in each set is 2-colored, by the pigeonhole principle, at least two of the first three elements in each set is monochromatic. Write S_a and S_{a+d} as follows;

$$S_a = \{s_{a;t},\ s_{a;t+m},\ s_{a;t+2m}, s_{a;t+3m},\ s_{a;t+4m}\},$$
$$S_{a+d} = \{s_{a+d;t},\ s_{a+d;t+m},\ s_{a+d;t+2m},\ s_{a+d;t+3m},\ s_{a+d;t+4m}\}.$$

Therefore, the two monochromatic elements in S_a become $s_{a;t}$ and $s_{a;t+m}$. To illustrate this notation, take $S_a = S_2 = [6, 10]$. Then $s_{2;1} = 6$, $s_{2;2} = 7$, $s_{2;3} = 8\, s_{2;4} = 9$, $s_{2;5} = 10$. The two elements, $s_{a;t}$ and $s_{a;t+m}$, belong to the first three elements in S_a therefore the difference between them is either 1 or 2 i.e.$m = 1$ or $m = 2$. Now $\{s_{a;t},\ s_{a;t+m},\ s_{a;t+2m}\}$ is an AP. If $s_{a;t+2m}$ is the same color as $s_{a;t}$ and $s_{a;t+m}$, then we have a monochromatic 3-term AP and we are done therefore assume that $s_{a;t+2m}$ is not the same color as $s_{a;t}$ and $s_{a;t+m}$. Since S_a and S_{a+d} are colored in the same way, $s_{a;t}$ and $s_{a+d;t+m}$ are monochromatic. Now we will consider the set $S_{a+2d} = \{s_{a+2d;t},\ s_{a+2d;t+m},\ s_{a+2d;t+2m},\ s_{a+2d;t+3m},\ s_{a+2d;t+4m}\}$. If $s_{a+2d;t+2m}$ is the same color as $s_{a;t}$ and $s_{a+d;t+m}$, we have a monochromatic 3-term AP and we are done therefore assume that $s_{a+2d;t+2m}$ is not the same color as $s_{a;t}$ and $s_{a;t+m}$. Since we are considering only 2-colorings of the elements, $s_{a+2d;t+2m}$ and $s_{a;t+2m}$ are monochromatic and since S_a and S_{a+d} are colored in the same way, $s_{a;t+2m}$, $s_{a+d;t+2m}$ and $s_{a+2d;t+2m}$ are monochromatic. Therefore, we have a monochromatic 3-term AP. Hence regardless of the way in which [1, 325] is 2-colored, a monochromatic 3-term AP exists and the result follows. □

The argument used in the above proof is called "*color-focusing*" and it is an example of the double induction technique performed on the number of colors r and the k number of terms in the AP's. The above proposition gives a very loose upper bound for $W(3, 2)$. The exact value of $W(3, 2)$ is 9, proven in [40].

Definition 4.4.1. Given an r-coloring of $\mathbb{Z}^+$ with the number of different k-term AP's being t. These k-term AP's are called *color-focused* if the following conditions hold;

(1) each k-term AP is monochromatic,

(2) consider the t different k-term AP's, none of them must be monochromatic in the same color and

(3) the $(k+1)$th terms of every AP is equal.

The $(k+1)$th term of the t color-focused AP's is called the *color-focus* of the AP's.

The following example elucidates this definition.

Example 4.4.1. Consider a 2-coloring of set $S = [1, 325]$ in red and green. Choose two monochromatic sets $S_{14} = \{\mathbf{61}, \mathbf{62}, \mathbf{63}, \mathbf{64}, \mathbf{65}\}$ and $S_{33} = \{\mathbf{161}, \mathbf{162}, \mathbf{163}, \mathbf{164}, \mathbf{165}\}$. These two sets do not yield any monochromatic 3-term AP. Take two different 2-colorings of the set S_{53}, namely $S_{53} = \{261, 262, \mathbf{263}, 264, 265\}$ and $S_{53} = \{261, 262, \mathbf{263}, 264, 265\}$, in which the black numbers are left uncolored and may be red or green. The first 2-coloring of S_{53} yields the monochromatic 3-term AP $\{\mathbf{63}, \mathbf{163}, \mathbf{263}\}$ and the second

yields {**61**, **162**, **263**}. The color-focus of the two monochromatic 2-term AP's is 263.

We now commence with the proof of van der Waerden's theorem, restated here for convenience.

Theorem 4.4.3. *Suppose k, $r \in \mathbb{Z}^+$, then there exists a least positive integer $W(k, r) = n$, such that any r-coloring of $[n]$ yields a monochromatic k-term AP.*

Proof. We first perform induction on k to show that $W(k, r) = n$ exists, by showing that it is bounded. Let $k = 1$. Then the statement is trivially true i.e. $W(1, r)$ exists. So suppose $k \geq 2$ and assume that for any $p \leq k$ and q, $W(p, q)$ exists (this is the induction hypothesis for k). We need to show that $W(k+1, r)$ exists for all r in order to complete the inductive argument. We will prove this using the following claim:

Claim:

For all $t \in [1, r]$, there exists a $W(t, k, r) = w$ such that whenever $[w]$ is r-colored, it yields either a monochromatic $(k+1)$-term AP in $[w]$, or t color-focused k-term AP's in $[w]$.

Proof of claim:

We perform induction on t. Let $t = 1$. By the induction hypothesis for k, there exists only one monochromatic k-term AP in $[n]$ therefore it must be color-focused. Call the color-focus of this AP, x (recall that x must be the $(k+1)$th term). Now $x \leq 2n$ therefore $[2n]$ also yields a color-focused k-term monochromatic AP. Hence $W(1, k, r) = 2n$ and the claim is true for $t = 1$, so let us assume that w exists. In order to complete the induction argument

we need to show that $W(t+1,\, k,\, r)$ exists. Firstly let $N = 2wW(k,\, r^w)$ and then break $\{1, \cdots, N\}$ into individual sets each of length w. Name each of these sets S_i where i denotes the position the set holds in $\{1, \cdots, N\}$ and since each $|S_i| = w$, we have $2W(k,\, r^w)$ S_i's in total.

$$\begin{aligned}\{1, \cdots, N\} &= \{1,\, 2, \cdots, w\} \cup \{w+1,\, w+2, \cdots, 2w)\} \cup \cdots \\ &\cup \{N-w+1,\, N-w+2, \cdots, N\} \\ &= S_1 \cup S_2 \cup \cdots \cup S_{2W(k,\, r^w)}.\end{aligned}$$

Now we r-color $[N]$, since $|S_i| = w$ in $[N]$, there are r^w different ways in which each S_i can be r-colored. Suppose any r-coloring of $[N]$ yields in any of the S_i's a monochromatic $(k+1)$-term AP, then we are done therefore assume the other condition of the claim is true i.e.each S_i yields a t color-focused k-term AP. Consider an r^w-coloring of $[W(k,\, r^w)]$, this must yield a monochromatic k-term AP. Any r-coloring of $[wW(k,\, r^w)]$ produces an r^w-coloring of the set $\{S_1,\, S_2, \cdots, S_{W(k,\, r^w)}\}$ since each $|S_i| = w$ and is therefore r-colored in one of r^w ways. Hence $\{S_1,\, S_2, \cdots, S_{W(k,\, r^w)}\}$ yields k monochromatic sets $S_a,\, S_{a+d}, \cdots, S_{a+(k-1)d}$ whose indices are an AP. Assume that each of these sets yields a t color-focused k-term AP with color-focus say x. Write each element of $\{1, \cdots, N\}$ as $s_{i;j}$, where i denotes which set S_i it belongs to and j denotes the position in that set. Label the t color-focused k-term AP's in

S_a as follows;

$$\begin{aligned}
A_{a;1} &= \{s_{a;u},\ s_{a;u+v},\ \cdots,\ s_{a;u+(k-1)v}\} \\
A_{a;2} &= \{s_{a;u_1},\ s_{a;u_1+v_1},\ \cdots,\ s_{a;u_1+(k-1)v_1}\} \\
&\vdots \\
A_{a;t} &= \{s_{a;u_t},\ s_{a;u_t+v_t},\ \cdots,\ s_{a;u_t+(k-1)v_t}\}.
\end{aligned}$$

The color-focus of these AP's being $s_{a;f}$, i.e. $s_{a;u+kv} = s_{a;u_1+kv_1} = s_{a;u_t+kv_t} = s_{a;f}$. S_a, S_{a+d}, $\cdots$ $S_{a+(k-1)d}$ produces monochromatic k-term A.P's (since they are colored in the same way). Name them as follows;

$$\begin{aligned}
A_{a;1} &= \{s_{a;u},\ s_{a;u+v},\ \cdots,\ s_{a;u+(k-1)v}\} \\
A_{a+d;1} &= \{s_{a+d;u},\ s_{a+d;u+v},\ \cdots,\ s_{a+d;u+(k-1)v}\} \\
&\vdots \\
A_{a+(k-1)d;1} &= \{s_{a+(k-1)d;u},\ s_{a+(k-1)d;u+v},\ \cdots,\ s_{a+(k-1)d;u+(k-1)v}\} \\
A_{a;2} &= \{s_{a;u_1},\ s_{a;u_1+v_1},\ \cdots,\ s_{a;u_1+(k-1)v_1}\} \\
&\vdots \\
A_{a+(k-1)d;2} &= \{s_{a+(k-1)d;u_1},\ s_{a+(k-1)d;u_1+v_1},\ \cdots,\ s_{a+(k-1)d;u_1+(k-1)v_1}\} \\
&\vdots \\
A_{a+(k-1)d;t} &= \{s_{a+(k-1)d;u_t},\ s_{a+(k-1)d;u_t+v_t},\ \cdots,\ s_{a+(k-1)d;u_t+(k-1)v_t}\}.
\end{aligned}$$

Let χ denote the coloring on the elements belonging to each AP, $A_{i;j}$. This

implies that

$$\begin{aligned}
\chi(A_{a;1}) = \chi(A_{a+d;1}) = \cdots = \chi(A_{a+(k-1)d;1}) \\
\chi(A_{a;2}) = \chi(A_{a+d;2}) = \cdots = \chi(A_{a+(k-1)d;2}) \\
\vdots \\
\chi(A_{a;t}) = \chi(A_{a+d;t}) = \cdots = \chi(A_{a+(k-1)d;t}).
\end{aligned}$$

There are t number of AP's in the k number of sets which yields a $t+1$ color-focused monochromatic k-term AP. Choose one particular k-term AP,

$$\begin{aligned}
P_1 &= \{s_{a;u},\ s_{a+d;u+v},\ \cdots,\ s_{a+(k-1)d;u+(k-1)v}\} \\
P_2 &= \{s_{a;u_1},\ s_{a+d;u_1+v_1},\ \cdots,\ s_{a+(k-1)d;u_1+(k-1)v_1}\} \\
&\vdots \\
P_t &= \{s_{a;u_t},\ s_{a+d;u_t+v_t},\ \cdots,\ s_{a+(k-1)d;u_t+(k-1)v_t}\}.
\end{aligned}$$

The elements in each P_i were picked out of $A_{j;i}$, where $j \in \{a,\ a+d,\ \cdots,\ a+(k-1)d\}$ and $i \in \{1,\ 2,\ \cdots,\ t\}$, therefore each P_i must be monochromatic and the $(k+1)$th term of each of these P_i's are equal i.e.

$$s_{a+kd;u+kv} = s_{a+kd;u_1+kv_1} = \cdots = s_{a+kd;u_t+kv_t} = s_{a+kd;f},$$

which is the color-focus of these the monochromatic k-term A.P's. The element $y \in N$ since $N = 2wW(k,\ r^w)$ (recall that all of these elements come from the first $wW(k,\ r^w)$ elements). Hence $s_{i;f}$, where $i = \{a,\ a+d,\ \cdots,\ a+kd\}$, must have the same color in every set of the monochromatic k-set AP.

Therefore the color-focuses of S_a, S_{a+d}, $\cdots$, $S_{a+(k-1)d}$ also form a monochromatic k-term AP, call this color-focus $s_{a+kd;f}$. Now consider the k-term AP, $s_{a;f}$, $s_{a+d;f}$, $\cdots$, $s_{a+(k-1)d;f}$. They must each have a different color in each of the $P_1, \cdots, P_t$ (from the color-focus definition). Hence $s_{a+kd;f}$ is the color-focus for the $t+1$ monochromatic k-term AP's therefore $N = W(t+1, k, r)$ and the claim has been proven by induction.

Since $w = W(t, k, r)$ exists $\forall t \in [1, r]$, by taking $t = r$ we have that $W(r, k, r)$ exists. This implies that in any r-coloring of $[W(r, k, r)]$, either an r color-focused k-term AP exists or a monochromatic $(k+1)$-term AP exists. If the second case is true, then we are done therefore assume that the first case is true. The color-focus of the r k-term AP's must have one of r colors therefore one of the r k-term AP's and this color-focus must have the same color. The color-focus combined with the r k-term AP forms a $(k+1)$-term AP. Hence by induction we have shown that for $k, r \in \mathbb{Z}^+$ then there exists a $W(k, r) = n \in \mathbb{Z}^+$, such that in any r-coloring of $[n]$, a monochromatic k-term AP exists.

□

4.5 Van Der Waerden Numbers

Definition 4.5.1. For $r, k_1, k_2, \cdots, k_r \in \mathbb{Z}^+$, a *generalised van der Waerden number*[4] $W(k_1, k_2, \cdots, k_r, r) = n \in \mathbb{Z}^+$, is the smallest n such that for every r-coloring of $[n]$ there exists a k_i-term AP monochromatic in color i.

[4]Also widely known as a *mixed van der Waerden number*

Van der Waerden numbers are also computionally tedious to find therefore few of them have actually been discovered. The table below presents a summary of a select few classical values, extracted from [21]. The most recent list of all known mixed van der Waerden numbers is found in [2] together with a thorough referencing of these values and bounds.

k \ r	**2**	**3**	**4**	**5**	**6**
3	9	27	76	> 125	> 207
4	35	> 292	> 1048	> 2254	> 9778
5	178	> 965	> 10437	> 24045	> 56693
6	> 1131	> 8886	> 90306	> 246956	> 600486
7	> 3703	> 43855	> 387967	-	-
8	> 7484	> 238400	-	-	-
9	> 27113	-	-	-	-

Table 4.2: Known values and bounds of van der Waerden numbers $W(k, r)$.

The history of results related to finding upper and lower bounds on van der Waerden numbers is entertaining, two achievements in this field in particular merit discussion. In 1987, renowned Israeli logician Saharon Shelah stunned the combinatorial community by discovering an ingenious new proof of the Hales-Jewett theorem and hence of van der Waerdens theorem, found in [35]. Shelah's proof is independent of van der Waerdens original double induction argument and gives rise to one of the most profound results on van der Waerden numbers. Before we state it however, some construction is

required.

Definition 4.5.2. For $k \geq 1$, define the following functions

(1) $t_1(k) = 2k$,

(2) $t_2(k) = t_1^{(k)}(1) = 2^k$, where $t^{(k)}$ is the composition of t with itself k times,

(3) $t_3(k) = t_2^{(k)}(1) = 2^{2^{\cdot^{\cdot^{\cdot^{2}}}}}$, where this expression contains k 2's and

(4) $t_{i+1}^{(k)} = t_i^{(k)}(1)$, for $i \in \mathbb{Z}^+$.

These functions clearly display rapid growth. We name $t_3(k) = Tower(k)$ and $t_4(k) = Wow(k)$, the $Tower$ and Wow functions, respectively and call $t_k(k) = Ack(k)$ the *Ackermann* function. This function is credited to Wilhelm Ackermann, a student of David Hilbert, who first derived it in [1]. There are many variations of his original function throughout mathematics and it is considered to be the fastest growing function one can encounter, here it has been modified for our limited purposes. The bound $W(k, 2) \leq Ack(k)$ can be gleaned from van der Waerdens original proof however Shelah improved this bound with the following theorem.

Theorem 4.5.1 (Shelah)**.** *Let $k \geq 2$, then*

$$W(k, 2) \leq Wow(k).$$

Proof. Refer to [35]. □

For many years, distinguished mathematician Ronald Graham had offered a \$1000 cash prize to whomever could prove $W(k,\,2) \leq Tower(k)$. Shelah was awarded half this prize money for the above result. The second more powerful result is attributed to Timothy Gowers, who won the Fields Medal in 1998 for work closely related to his paper [15], in which he developed Shelah's bound even further. The theorem below played a considerable role in him achieving this honour.

Theorem 4.5.2 (Gowers)**.** *Let $k \geq 2$, then*

$$W(k,\,2) \leq 2^{2^{2^{2^{2^{k+9}}}}}.$$

Proof. Refer to [15]. □

Shelah and Gowers both resolved upper bounds for van der Waerden numbers, the only lower bound of note was constructed in [10] by Elwyn Berlekamp in 1968.

Theorem 4.5.3 (Berlekamp)**.** *For prime p,*

$$W(p+1,\,2) \geq p2^p.$$

Proof. Refer to [10]. □

4.6 Schur's Theorem

Issai Schur's influential paper [34] was inspired by Fermat's last theorem. In his original paper he proved that for all m and for suffiently large primes p, the equation $x^m + y^m = z^m (\mathrm{mod}\, p)$, has a non-trivial solution.

Theorem 4.6.1 (Schur)**.** *For $r \in \mathbb{Z}^+$, there exists a least positive integer $S(r) = s$, such that for any r-coloring of $[s]$, there exists a monochromatic solution to the equation $x + y = z$.*

Proof. The proof is an application of Ramsey's theorem (Theorem 4.2.1) and is more or less identical to Schur's original proof. Theorem 4.2.1 asserts that for an integer $R(3, r) = n$, any r-coloring of K_n yields a monochromatic triangle K_3. We therefore define a particular coloring of K_n as follows;

(1) label the vertices of K_n using the set $\{1, 2, \cdots, n\}$,

(2) then perform an r-coloring on the set of vertices $\{1, 2, \cdots, n-1\}$, i.e.color each vertex $v \in \{1, 2, \cdots, n-1\}$ in one of the r colors,

(3) now perform an edge-coloring on K_n, by taking 2 vertices labeled i and j, $i > j$, calling this edge ij and lastly coloring ij and $i - j$ with the same color.

By Theorem 4.2.1, K_n yields a monochromatic triangle K_3. Label the vertices of this triangle a_1, a_2, a_3, where $a_1 < a_2 < a_3$, then $a_2 - a_1$, $a_3 - a_2$ and $a_3 - a_1$ all have the same color. Take $x = a_2 - a_1$, $y = a_3 - a_2$ and $z = a_3 - a_1$, then

$$x + y = (a_2 - a_1) + (a_3 - a_2) = a_3 - a_1 = z$$

and the proof is complete. □

The set $\{x, y, z\}$ that satisfies $x + y = z$ is called a monochromatic *Schur triple* and $S(r) \in \mathbb{Z}^+$ are called *Schur numbers*. This definition is formalised below.

Definition 4.6.1. For $r \in \mathbb{Z}^+$, a *Schur number* $S(r) = s \in \mathbb{Z}^+$ is the smallest s, such that any r-coloring of $[s]$ yields a monochromatic solution to the equation $x + y = z$.

Schur numbers are even more vexing than Ramsey numbers and van der Waerden numbers to calculate, the table below adapted from [13] delivers some familiar values whilst [3] tabulates all *generalised Schur numbers* that have been found to date.

$\boldsymbol{r}$	**1**	**2**	**3**	**4**	**5**	**6**	**7**
$\boldsymbol{S(r)}$	1	5	14	45	[160, 305]	[536, ∞)	[1680, ∞)

Table 4.3: Known values and bounds for Schur numbers $S(r)$.

We conclude this section by generalising Theorem 4.6.1, another more compelling generalisation of the same result is referred to as Rado's theorem, which we will examine in the next section.

Theorem 4.6.2 (Generalised Schur's Theorem)**.** *Let $r \in \mathbb{Z}^+$ and select a color $i \in [1, r]$. There exists a least positive integer $S(k_1, k_2, \cdots, k_r) = s$, where $k_i \geq 3$, such that every r-coloring of $[s]$ yields a solution to the equation*

$$x_1 + x_2 + \cdots + x_{k_j - 1} = x_{k_j} \tag{4.1}$$

that is j-monochrome for color $j \in [1, r]$.

Proof. Similar to the proof of Schur's theorem, this proof is an application of the generalised Ramsey theorem (Theorem 4.2.2). Let the generalised Ramsey number $R(k_1, k_2, \cdots, k_r) = n$ and consider an r-coloring of $[n]$.

We utilise the same edge-coloring as in the proof of Theorem 4.6.1 (Schur's theorem). By Theorem 4.2.2, this edge-coloring of K_n must yield a K_{k_j} subgraph that is j-monochrome in some color $j \in [1, r]$. Let $a = k_j$ and denote the vertices of K_a by $\{v_0, v_1, \cdots, v_{a-1}\}$ and define a difference $d_i = v_i - v_0$ and assume that $d_{i-1} < d_i$, $\forall i \in [2, a-1]$. Now K_a is j-monochrome therefore the edges $e = v_{i-1}v_i$ and $e' = v_{a-1}v_0$ for $i \in [1, a-1]$, are also j-monochrome. Consider

$$v_i - v_{i-1} = (d_i + v_0) - (d_{i-1} + v_0) = d_i - d_{i-1},\ \forall i \in [2, a-1],$$

therefore $d_i - d_{i-1}$, d_1 and d_{a-1} are all j-monochrome, $\forall i \in [2, a-1]$. Hence

$$d_1 + \sum_{i=2}^{a-1} (d_i - d_{i-1}) = d_{a-1} \Longrightarrow d_1 + d_2 + \cdots + d_{i-2} = d_{i-1}$$

for any r-coloring of $[n-1]$.

□

4.7 Rado's Theorem

The generalised Schur's theorem declares that in any r-coloring of $\mathbb{Z}^+$, monochromatic solutions to equations of the form $\sum_{i=1}^{k-1} x_i = x_k$ exist for $k \geq 4$, this assertion can be extended to include systems of *homogeneous linear equations* over $\mathbb{Z}$ i.e. equations of the form $\sum_{i=1}^{k} a_i x_i = 0$, where the coefficients $0 \neq a_i \in \mathbb{Z}$.

Definition 4.7.1. Let $\mathfrak{L}$ be a system of linear homogeneous equations and $r \geq 1$. If there exists a monochromatic solution to $\mathfrak{L}$, for every r-coloring of $\mathbb{Z}^+$, then $\mathfrak{L}$ is said to be *r-regular*. $\mathfrak{L}$ is said to be just *regular* if it is r-regular $\forall r \geq 1$.

Schur's theorem shows that the equation $x + y = z$ is regular, this result was strengthened by Rado in his 1933 thesis [29]. Here we prove regularity for single homogeneous linear equations. For a system of such equations Rado determined a necessary and sufficient condition called the "*Columns Condition*" that must be satisfied for regularity to be achieved, confer with [17] and [29] for a complete proof of this.

Theorem 4.7.1 (Rado's Single Equation Theorem)**.** *Let $k \geq 2$ and $0 \neq a_i \in \mathbb{Z}$, for $i \in [1, k]$. The equation*

$$\sum_{i=1}^{k} a_i x_i = 0 \tag{4.2}$$

is regular $\iff$ there exists a nonempty set $B \subseteq \{a_i : i \in [1, k]\}$ such that $\sum_{b \in B} b = 0$.

We detour to the following example to clarify the statement of this theorem before proving it.

Example 4.7.1. The statement effectively says that equation (4.2) is regular $\iff$ there is some nonempty subset of the coefficients a_i that add up to 0. Hence the equation $x_1 + 2x_2 - 3x_3 = 0$ is regular whereas $2x_1 + 3x_2 - 4x_3 = 0$ is not regular.

Proof. We will first show the "$\Longleftarrow$" statement i.e. we assume, $\forall r \geq 1$, there exists a set $B \neq \emptyset$ defined as above with $\sum_{b\in B} b = 0$, then we will show that (4.2) is r-regular. We will prove this by induction on r but first w.l.o.g take $B = \{a_1, a_2, \cdots, a_m\}$, where $a_1 > 0$ and $|B| = m$ is maximal. Suppose $m = k$, then

$$\sum_{i=1}^{k} a_i = \sum_{b\in B} b = 0.$$

Therefore we can take $x_i = 1$, for $i \in [1, k]$ to be our monochromatic solution since

$$\sum_{i=1}^{k} a_i x_i = \sum_{i=1}^{k} a_i = 0.$$

Hence we let $m < k \Longrightarrow c = \sum_{i=m+1}^{k} a_i$ is nonempty. Now take

$$x_2 = x_3 = \cdots = x_m \text{ and } x_{m+1} = x_{m+2} = \cdots = x_k. \tag{4.3}$$

Then (4.2) may be written as

$$a_1x_1 + (a_2 + a_3 + \cdots + a_m)x_2 + (a_{m+1} + a_{m+2} + \cdots + a_k)x_{m+1} = 0,$$

since $a_1 + a_2 + \cdots + a_m = 0$, (4.2) reduces to

$$a_1(x_1 - x_2) + cx_{m+1} = 0. \tag{4.4}$$

We now begin the induction process.

Let $r = 1$:

For $x_1, x_2 \in \mathbb{Z}^+$, if we take $c = x_2 - x_1$ and $x_{m+1} = a_1$, then we have a

monochromatic solution to (4.4).

Let $r \geq 2$:

Suppose the "$\Longleftarrow$" statement is true, $\forall t \in [1,\, r-1]$, we will show that it is true for all r. For all $t \in [1,\, r-1]$ define $\mu(t) \in \mathbb{Z}^+$ to be smallest $\mu(t)$ such that for any t-coloring of $[\mu(t)]$, there exists a monochromatic solution to (4.2). Clearly the induction hypothesis shows that $\mu(t)$ exists. Suppose w.l.o.g that $\sum_{i=1}^{m} a_i = 0$, for $a_1 > 0$ and m maximal, $m \neq k$ therefore $c \neq 0$. Take $d = \sum_{i=1}^{k} |a_i|$ and $\mu = \mu(r-1)$. We will show that $\mu(r) \leq dW(\mu+1,\, r)$, where $W(\mu+1,\, r)$ denotes the van der Waerden number. We need to show that for any r-coloring of $[dW(\mu+1,\, r)]$, a monochromatic solution to (4.2) exists. Similar to the the case $r = 1$, assume (4.3) hence equation (4.2) becomes (4.4). Suppose we have an r-coloring χ of $[dW(\mu+1,\, r)]$, we will proceed to find a monochromatic solution to (4.4) i.e. x_1, x_2 and x_{m+1} such that $\chi(x_1) = \chi(x_2) = \chi(x_{m+1})$. Remember that $0 \neq c = \sum_{i=m+1}^{k} a_i$ and $d = \sum_{i=1}^{k} |a_i| \implies 1 \leq |c| < d$. Now choose an $f \in [1,\, d]$ and let χ_f denote an r-coloring of $[W(\mu+1,\, r)]$ such that $\chi_f(i) = \chi(fi)$, $\forall i \in [W(\mu+1,\, r)]$. This implies that, $\forall f \in [1,\, d]$, there exists a monochromatic set which is monochromatic w.r.t χ_f (follows from Theorem 4.4.3), namely

$$A = \{\alpha,\, \alpha+q,\, \alpha+2q,\, \cdots,\, \alpha+\mu q\}.$$

This implies that, $\forall y \in A$, $y \leq W(\mu+1,\, r)$. Therefore

$$\chi_f(\alpha) = \chi(f\alpha) = \chi_f(\alpha+q) = \chi(f\alpha+fq) = \cdots = \chi_f(\alpha+\mu q) = \chi(f\alpha+f\mu q)$$

i.e. the set

$$fA = \{f\alpha,\, f\alpha + fq,\, f\alpha + f2q,\, \cdots,\, f\alpha + f\mu q\}$$

is monochromatic with respect to χ and $\forall y \in fA,\ y \leq fW(\mu+1,\, r) \implies y \in [dW(\mu+1,\, r)]$. Now take $f = |c|$ and $e = f\alpha$ therefore, $\forall q \geq 1$, there exists a monochromatic set (from Theorem 4.4.3),

$$\{e,\, e + q|c|,\, e + 2q|c|,\, \cdots,\, e + \mu q|c|\}$$

such that $e + \mu q|c| \leq dW(\mu+1,\, r)$. Consider the following cases.

Case 1:

If $\exists j \in [\mu]$ such that $\chi(jqa_1) = \chi(e)$ and if we let $|x_2 - x_1| = jq|c|$ and $x_{m+1} = jqa_1$, then we have a monochromatic solution to (4.4).

Case 2:

If $\exists j \in [\mu]$ such that $\chi(jqa_1) \neq \chi(e)$, then the set

$$\{qa_1,\, 2qa_1,\, \cdots,\, \mu qa_1\} = qa_1[\mu]$$

is $(r-1)$-colored and so we have a monochromatic solution to (4.4). Hence we have shown the "$\Longleftarrow$" statement. We will now prove the "$\Longrightarrow$" statement i.e. we assume that $\sum_{i=1}^{k} a_i x_i = 0$ is regular, then we need to show that there exists a nonempty set $B \subseteq \{a_i : i \in [1,\, k]\}$ such that $\sum_{b \in B} b = 0$. We will make a contrapositive argument, therefore assume that there exists $a_1,\, a_2,\, \cdots,\, a_k$ such that the elements of no nonempty subset of B sums to 0 and we need to show that, for $r \geq 1$, there exists an r-coloring on $\mathbb{Z}^+$ with no monochromatic

solution to (4.2). We begin by selecting a prime p such that for set $C \subseteq \{a_i : i \in [1, k]\}$, $p \nmid \sum_{c \in C} c$ (p does not divide $\sum_{c \in C} c$). Now C is a finite set so this is feasible. Take $r = p - 1$ and define a $(p-1)$-coloring, $\chi : \mathbb{Z}^+ \longrightarrow [p-1]$ as follows; $\forall i \in \mathbb{Z}^+$, take c to be the largest value such that $p^c \mid i$, for $i \in \mathbb{Z}^+ \implies \exists j \not\equiv 0 \pmod{p}$ such that $i = p^c j$. Then define $\chi(i) = j \equiv 0 \pmod{p}$. We need to show that (4.2) doesn't have a monochromatic solution with respect to χ, therefore we assume that one such solution exists and proceed to find a contradiction. So call this solution $X = \{x_1, x_2, \cdots, x_k\}$, that is v-monochrome w.r.t χ for color $v \in [1, p-1]$. Then, $\forall x_i \in X$, $\exists c_i, k_i$ such that $x_i = p^{c_i}(pk_i + v)$. Now $c = \min \{c_1, c_2, \cdots, c_k\} \implies$

$$\sum_{i=1}^{k} a_i x_i = p^c \sum_{i=1}^{k} c + i p^{c_i - r}(pk_i + v) = 0$$

$\implies$

$$\sum_{i=1}^{k} a_i p^{t_i}(pk_i + v) = 0$$

where $t_i = c_i - c \geq 0, \forall i$, hence

$$0 \equiv v \sum_{i=1}^{k} p^{t_i} a_i \pmod{p}.$$

Now since p is prime and $p \nmid v$ and taking $t_i = 0 \implies p \mid \sum_{i=1}^{k} a_i$. Let $C = \{a_i : i \in [1, k], t_i = 0\} \neq \emptyset$ but $p \nmid \sum_{a_i \in C} a_i$, this is a contradiction hence the "$\implies$" statement follows. $\square$

4.8 Hales-Jewett Theorem

We conclude our discourse of Ramsey theory with another fundamental result called the Hales-Jewett theorem. In 1963, the mathematicians Alfred W. Hales and Robert I. Jewett, revised van der Waerden's double induction proof in [20] to search for monochromatic *combinatorial lines* in arbitrary r-colorings of k-dimensional hypercubes $[n]^k$. This quote from [17] perfectly captures the significance of the result,

> "*The Hales-Jewett theorem strips van der Waerden's theorem of its unessential elements and reveals the heart of Ramsey theory. It provides a focal point from which many results can be derived and acts as a cornerstone for much of the more advanced work.*"

Only two combinatorial proofs of the Hales-Jewett theorem exist, the original proof given in [20] and Shelah's aforementioned proof found in [35]. Stating the Hales-Jewett theorem first requires some terminology.

Definition 4.8.1. Let $k, n \in \mathbb{Z}^+$. Then $[n]^k = [n] \times [n] \times \cdots \times [n]$ (k times) denotes the set of all k-tuples on $[n]$ i.e. sequences of length k whose elements belong to $[n]$. Equivalently, $[n]^k$ may be pictured as a k-dimensional cube whose sides are of length n.

Definition 4.8.2. Take an element $y \in [n]^k$, then $y = (y_1, y_2, \cdots, y_k)$ where $y_i \in [n], \forall i$. The element y is called a *word* of length n written as $y = y_1 y_2 \cdots y_n$.

Definition 4.8.3. A *variable word* is a word $y \in ([n] \cup \{\star\})^k$, where $k \geq 1$ and at least one $y_i = \star$ for a fixed value $\star \in [n]$.

Definition 4.8.4. Let l be a variable word and define a word $l(i)$, $i \in [n]$, which is derived by replacing all occurences of $\star$ in l by i.

Definition 4.8.5. A combinatorial line is a set $L = \{l(i) : i \in [n], l \text{ is a variable word}\}$.

Example 4.8.1. Take $n = 5$ and $k = 5$ and let variable word $l = 12 \star 3\star$. Then l generates the combinatorial line

$$L = \{12\mathbf{1}3\mathbf{1}, 12\mathbf{2}3\mathbf{2}, 12\mathbf{3}3\mathbf{3}, 12\mathbf{4}3\mathbf{4}, 12\mathbf{5}3\mathbf{5}\}.$$

Definition 4.8.6. Let $\chi : [n]^k \longrightarrow [r]$ be a coloring. A combinatorial line L is said to be *monochromatic* w.r.t χ if $\chi(l(1)) = \chi(l(2)) = \cdots = \chi(l(n))$.

Theorem 4.8.1 (Hales-Jewett)**.** *For k, r, $n \in \mathbb{Z}^+$, there exists a $HJ(n, r) = h \in \mathbb{Z}^+$ such that $\forall k \geq h$, every r-coloring of $[n]^k$ yields a monochromatic combinatorial line.*

Proof. Refer to [16] or [35]. □

The Hales-Jewett theorem affords us a much simpler way of proving van der Waerden's theorem (Theorem 4.4.3) as opposed to the naive proof given earlier.

The Hales-Jewett theorem implies van der Waerden's theorem:

Proof. Let $HJ(k, r) = n$ and suppose that $n = n'(k-1) + 1$. Now define a map $f : [k]^n \longrightarrow [n]$, where word $y = (y_1, y_2, \cdots, y_n) \in [k]^n$ and $f(y) = y_1, y_2, \cdots, y_n$. We then use f to define an r-coloring $\chi : [k]^n \longrightarrow [r]$, where $\chi(y) = \chi(f(y))$, for $y \in [k]^n$. This implies that f maps every combinatorial

line $L = \{l(1), l(2), \cdots, l(k)\}$ into a k-term AP. The difference between any two variable words in this k-term AP, $l(i)$ and $l(i+1)$, is the same and is equal to the number of $\star$'s in l. Therefore by the Hales-Jewett theorem, there exists a monochromatic combinatorial line that yields a monochromatic k-term AP and hence the statement of van der Waerden's theorem follows. $\square$

Chapter 5

Symmetric Colorings

This chapter concentrates on the fundamental results associated with colorings related to symmetry. Several formulas for enumerating symmetric colorings are made manifest from a group theoretic perspective. The results contained in these chapters are modeled from the authoritative papers [19], [43], [44], [45], [46] and [47]. Please note that all groups considered henceforth are finite.

Definition 5.0.7. Let $r \in \mathbb{N}$, an r-coloring of a group G is a mapping $\chi : G \longrightarrow [r]$. The set of all r-colorings of G is denoted by r^G.

Proposition 5.0.2. *Let G be a group and $r \in \mathbb{N}$. Then $r^{|G|}$ is the total number of all r-colorings of G.*

Proof. Let $\chi : G \longrightarrow [r]$, then there are r different ways of assigning one of the r colors to each of the elements of G, therefore the total number of different ways of coloring G with r colors is, $r \times r \times \cdots \times r$, $|G|$ times. Hence the number of r-colorings of G equals $r^{|G|}$. □

Definition 5.0.8. Let G be a group with the set $X \subseteq G$. We call X *symmetric* w.r.t an element $g \in G \iff gX^{-1}g = G$.

Definition 5.0.9. Let G be a group and take an element $g \in G$. The symmetry on G w.r.t an element g is a map $f_g : G \longrightarrow G$ defined as follows, for x, $gx^{-1}g \in G$,

$$f_g(x) = gx^{-1}g.$$

Definition 5.0.10. A coloring χ of a group G is called *symmetric* if there exists an element $g \in G$ such that $\forall x \in G$, $\chi(gx^{-1}g) = \chi(x)$. Let $r \in \mathbb{N}$. Then the set of all symmetric r-colorings of G is denoted by $S_r(G)$.

Definition 5.0.11. For $r \in \mathbb{N}$ and group G, let $\chi : G \longrightarrow [r]$. Now define another coloring χg as follows, for g, $x \in G$,

$$\chi g(x) = \chi(xg^{-1}).$$

We may deduce from the above definition that G *acts* (by means of a *right group action*) on r^G.

Definition 5.0.12. Consider two colorings of a group G, namely χ and θ. These two colorings are said to be *equivalent* if there exists an element $g \in G$ such that $\forall x \in G$, $\chi(xg^{-1}) = \theta(x)$.

Equivalent, in the above definition, specifies an *equivalence relation* on r^G, denoted by $\sim$.

Definition 5.0.13. For $r \in \mathbb{N}$ and group G, let coloring $\chi : G \longrightarrow [r]$. We write $[\chi]$ to denote the *orbit* of coloring χ and $St(\chi)$ to denote the *stabilizer* of χ i.e.

$$[\chi] = \{\chi g \,:\, g \in G\} \text{ and } St(\chi) = \{g \in G \,:\, \chi g = \chi\}.$$

Therefore $\sim$ also relates the *partitioning* of the colorings into orbits i.e.

$$\chi \sim \theta \iff \exists\ g \in G \text{ such that } \chi(xg^{-1}) = \theta(x)\,, \forall x \in G.$$

Definition 5.0.14. Each equivalence class of the r-colorings of group G is called a *necklace*. The number of all r-ary necklaces of G is denoted by $N_r(G)$.

Theorem 5.0.3 (Necklace Counting Theorem Version 1)**.** *Let $r \in \mathbb{N}$ and suppose the group G acts on r^G. Then*

$$N_r(G) = \frac{1}{|G|}\sum_{g\in G} r^{|G:\langle g\rangle|}. \tag{5.1}$$

Proof. The proof utilizes Burnside's lemma (Lemma 2.5.2). Let $\chi \in r^G$, since G acts on r^G, there exists a map $f : r^G \times G \longrightarrow r^G$ such that $f(\chi(x),\, g) = \chi g(x) = \chi(xg^{-1})$, for $x,\, g \in G$. Therefore the element x is translated into xg^{-1} (by f), similarly xg^{-1} is translated into xg^{-2}. Continuing in this fashion we can deduce that xg^{-j} is translated into $xg^{-(j+1)}$. Now $|\langle g\rangle| = o(g)$, therefore the elements of G are permuted into $|G : \langle g\rangle|$ different cycles and since each cycle is r-colored, the number of colorings in r^G fixed by g is $r^{|G:\langle g\rangle|}$. Hence, by Burnsides lemma,

$$N_r(G) = \frac{1}{|G|}\sum_{g\in G} r^{|G:\langle g\rangle|}.$$

□

Proposition 5.0.4. *Any r-coloring which is equivalent to a symmetric r-coloring is symmetric.*

Proof. Let χ, $\chi a \in r^G$, for $a \in G$ and let χ and χa be equivalent. Now χa is symmetric w.r.t $g \in G$, $\forall a \in G$. We are required to show that χ is also symmetric w.r.t some element in G. Now, for $x \in G$,

$$\begin{aligned} \chi a(gx^{-1}g) &= \chi a(x) \\ \chi a(x) &= \chi(xa^{-1}) \\ \chi(ga^{-1}(xa^{-1})^{-1}ga^{-1}) &= \chi(xa^{-1}) \\ \chi(ga^{-1}x^{-1}ga^{-1}) &= \chi(x). \end{aligned}$$

Therefore χ is symmetric w.r.t $ga^{-1} \in G$.

□

The set of all symmetric necklaces is denoted by $S_r(G)/\sim$. The following theorem gives an archetypal formula for computing $|S_r(G)|$ and $|S_r(G)/\sim|$.

Theorem 5.0.5. *Let G be an abelian group, then*

$$|S_r(G)| = \sum_{X \leq G} \sum_{Y \leq G} \frac{\mu(Y, X)|G/Y|}{|B(G/Y)|} r^{(|G/X|+|B(G/X)|)/2} \tag{5.2}$$

where $B(G)$ denotes the Boolean group, $\leq$ denotes the subgroup relation and

$$|S_r(G)/\sim| = \sum_{X \leq G} \sum_{Y \leq G} \frac{\mu(Y, X)}{|B(G/Y)|} r^{(|G/X|+|B(G/X)|)/2} \tag{5.3}$$

where $\mu(Y, X)$ denotes the Möbius function on the lattice of subgroups of G.

Proof. See [19]. □

This theorem is then generalised in [19] to include arbitrary groups. We must first construct the poset of optimal partitions of a group G to accomplish this. In the case of an abelian group G, the poset of optimal partitions coincides with the subgroup lattice of G.

Theorem 5.0.6. *Let P denote the poset of optimal partitions of a group G, then*

$$|S_r(G)| = |G| \sum_{x \in P} \sum_{y \leq x} \frac{\mu(y,\, x)}{|Z(y)|} r^{|x|} \tag{5.4}$$

$$|S_r(G)/\sim| = \sum_{x \in P} \sum_{y \leq x} \frac{\mu(y,\, x)|St(y)|}{|Z(y)|} r^{|x|} \tag{5.5}$$

where $\mu(x,\, y)$ denotes the Möbius function on P.

Proof. See [19]. □

A particularly interesting case of symmetric colorings of groups is when we take a finite group $G = \mathbb{Z}_n$.

Definition 5.0.15. A *regular n-gon* is a regular polygon with n sides (for example a pentagon is a 5-gon). Furthermore, all sides and angles are equal in a regular n-gon.

We may revise the definitions given previously to affect only a finite abelian group G. Now G *acts* on r^G (by means of a *right group action*) as follows, for $g, x \in G$,

$$(g + \chi)(x) = \chi(x - g).$$

Definition 5.0.16. Two colorings χ and θ of G are called *equivalent* if $\exists g \in G$ such that $\chi(x - g) = \theta(x)\,, \forall x \in G$.

Definition 5.0.17. A *(proper) symmetry* of G is a mapping $f_g : G \longrightarrow G$ such that

$$f_g(x) = g - x,$$

which is the same as

$$f_g(x) = 2g - x.$$

Definition 5.0.18. Let $\chi \in r^G$. Then χ is called *symmetric* if $\exists g \in G$ such that for g, $x \in G$,

$$\chi(g - x) = \chi(x),$$

which may be alternatively written as

$$\chi(2g - x) = \chi(x).$$

The group $\mathbb{Z}_n$ can be expressed geometrically as the vertices of a regular n-gon where an r-coloring is performed on the vertices of this n-gon. Then the (proper) symmetries of $\mathbb{Z}_n$ become the reflections of the n-gon about an axis drawn connecting two of the vertices, passing through the center of the n-gon. Let $\chi \in r^{\mathbb{Z}_n}$, then χ is called symmetric if it is invariant under any reflection of the n-gon. Suppose we have χ, $\theta \in r^{\mathbb{Z}_n}$, these two colorings are called equivalent if χ may be obtained from θ or θ from χ by any rotation of the n-gon. We will now impose respective formulas that enumerate the

colorings of $\mathbb{Z}_n$. The following expression for the number of necklaces of $\mathbb{Z}_n$ was first derived by Percy MacMahon in [27], hence it is commonly known as *MacMahon's formula*.

Theorem 5.0.7 (Necklace Counting Theorem Version 2)**.** *Let n, $r \in \mathbb{Z}^+$. The number of necklaces of $\mathbb{Z}_n$ is*

$$N_r(n) = \left(\frac{1}{n}\right) \sum_{d|n} \phi(d) r^{\left(\frac{n}{d}\right)}$$

where ϕ denotes the Euler function and $d|n$ means that d divides n.

Proof. Consider $r^{\mathbb{Z}_n}$, we need to ascertain how many r-colored necklaces of length n exist. Permuting the vertices of an r-colored necklace by moving the vertex at position i to position $(i+1)$ does not change the necklace in any way. If the n vertices are permuted in this way, then eventually they must come back to the initial r-coloring. Suppose this occurs after say d permutations, where $d|n$. The smallest number of permutations for a come-back is called the *period* of the necklace. For example, consider an 3-coloring of a 6-gon in the colors red (R), blue (B) and green (G). Write the 3-colored vertices in a *string* and permute them as follows;

$$\text{RGBRGB} \to \text{BRGBRG} \to \text{GBRGBR} \to \text{RGBRGB}.$$

Therefore the string RGBRGB has period 3. Now suppose we have an r-colored string of length n and period d. Counting itself, it has d permutations before yielding the original r-colored necklace. This r-colored necklace is unique, since

$$N_r(n) = \sum_{d|n} \left(\frac{1}{d}\right) S_d,$$

where S_d denotes the number of strings of period d. It is easily deducible that

$$S_n = \sum_{d|n} S_d,$$

where S_n denotes the number of strings of length n, but $S_n = r^n$ and so

$$r^n = \sum_{d|n} S_d,$$

by applying Möbius inversion (Theorem 3.4.1)

$$S_d = \sum_{x|n} \mu\left(\frac{d}{x}\right) r^x.$$

Hence

$$\begin{aligned} N_r(n) &= \sum_{d|n} \left(\frac{1}{d}\right) \sum_{x|n} \mu\left(\frac{d}{x}\right) r^x \\ &= \left(\frac{1}{n}\right) \sum_{d|n} \phi(d) r^{\left(\frac{n}{d}\right)}. \end{aligned}$$

The simplification on the right uses the formula $\phi(n) = \sum_{d|n} \mu\left(\frac{n}{d}\right) d$, which is proven in [9]. □

The following theorem calculates the symmetric colorings of cyclic groups of order n, denoted by C_n.

Theorem 5.0.8. *Whenever p is prime, d is a divisor of n, m is odd and $l \geq 1$*

$$|S_r(C_n)| = \begin{cases} \sum_{d|n} d \prod_{p|\frac{n}{d}} (1-p) r^{\frac{d+1}{2}}, & \text{for } n \text{ odd} \\ \sum_{d|\frac{n}{2}} d \prod_{p|\frac{n}{2d}} (1-p) r^{d+1}, & \text{for } n = 2^l m \end{cases} \tag{5.6}$$

and

$$|S_r(C_n)/\sim| = \begin{cases} r^{\frac{n+1}{2}}, & \text{for } n \text{ odd} \\ \frac{1}{2}\left(r^{\frac{n}{2}+1} + r^{\frac{m+1}{2}}\right), & \text{for } n = 2^l m \end{cases} \tag{5.7}$$

Proof. Formula (5.6) is proven in [43] and (5.7) is proven in [44]. □

The ensuing sections will enumerate the symmetric colorings of select well-known groups of small order, to demonstrate formulas (5.2) to (5.7). We first draw the lattice of subgroups $L(G)$ then we construct the poset of optimal partitions, by employing Proposition 3.3.1.

5.1 Symmetric Colorings of the Klein four group V_4

The total number of r-colorings of V_4 is r^4 (observed from Proposition 5.0.2). Recall that $V_4 = \langle a, b : a^2 = b^2 = \mathbf{e}, ab = ba \rangle$, $V_4 \cong D_2$ and $V_4 \cong C_2 \times C_2$. The lattice of subgroups of V_4, $L(V_4)$ is drawn below.

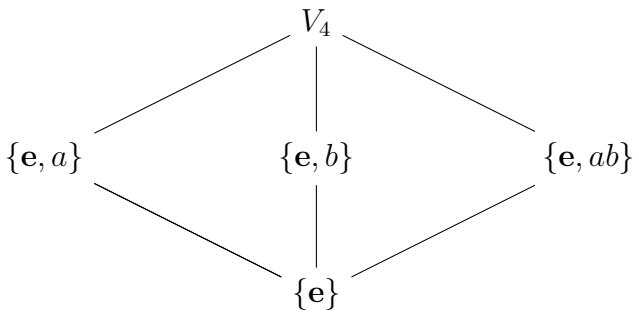

Figure 5.1: Lattice of subgroups of V_4

Since V_4 is abelian we may utilise Theorem 5.0.5,

$$\begin{aligned} |S_r(V_4)| &= r^4 + 3r^2(-1+1) + r^2(2-1-1-1+1) \\ &= r^4, \end{aligned}$$

$$\begin{aligned} |S_r(V_4)/\sim| &= \frac{1}{4}r^4 + 3r^2(-\frac{1}{4}+\frac{1}{2}) + r^2(\frac{2}{4}-\frac{1}{2}-\frac{1}{2}-\frac{1}{2}+1) \\ &= \frac{1}{4}r^4 + \frac{3}{4}r^2 \\ &= \frac{1}{4}r^2(r^2+3). \end{aligned}$$

Hence we have the following proposition.

Proposition 5.1.1. *For $r \in \mathbb{N}$,*

$$|S_r(V_4)| = r^4 \tag{5.8}$$

and

$$|S_r(V_4)/\sim| = \frac{1}{4}r^2(r^2+3), \tag{5.9}$$

therefore all the r-colorings of V_4 are symmetric.

Example 5.1.1. Let us take $r = 2$, then the number of symmetric 2-colorings of V_4, $|S_2(V_4)| = 2^4 = 16$ and $|S_2(V_4)/\sim| = 7$.

5.2 Symmetric Colorings of the Dihedral group D_3

The total number of r-colorings of D_3 is r^6. Formulas for the number of symmetric r-colorings of D_3 will be derived in this section, appropriated from material in [24]. Recall that $D_3 = \langle a, b : a^3 = \mathbf{e}, b^2 = \mathbf{e}, ba = a^{-1}b\rangle$, $D_3 \cong C_3 \ltimes C_2$ and $S_3 \cong D_3$, this group has 4 nontrivial subgroups depicted in $L(D_3)$, drawn below.

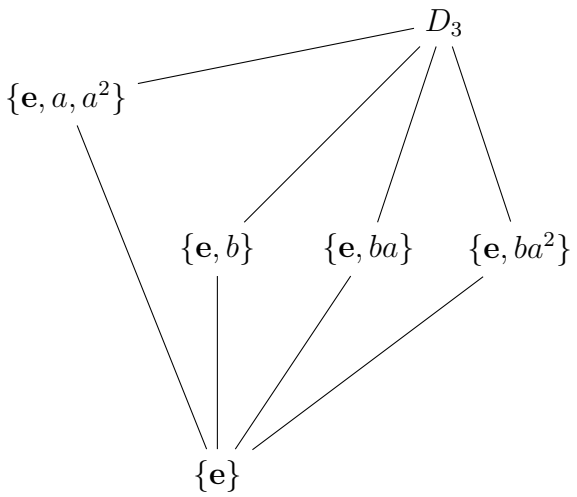

Figure 5.2: Lattice of subgroups of D_3

The optimal partitions of D_3 together with their respective stabilizers and centralizers as well as their orders are evaluated below.

(1) The finest partition

$\pi = \{\{\mathbf{e}\}, \{b\}, \{ba\}, \{ba^2\}, \{a, a^2\}\}$,

$St(\pi) = \{\mathbf{e}\}$, $Z(\pi) = \{\mathbf{e}\}$,

$|St(\pi)| = 1$, $|Z(\pi)| = 1$ and $|\pi| = 5$.

(2) Three partitions of the form

$\pi = \{\{\mathbf{e}\}, \{a, a^2\}, \{b, ba\}, \{ba^2\}\}$,

$St(\pi) = \{\mathbf{e}\}$, $Z(\pi) = \{\mathbf{e}, ba^2\}$,

$|St(\pi)| = 1$, $|Z(\pi)| = 2$ and $|\pi| = 4$.

(3) One partition

$\pi = \{\{\mathbf{e}, a, a^2\}, \{b\}, \{ba\}, \{ba^2\}\}$,

$St(\pi) = \{\mathbf{e}\}$, $Z(\pi) = \{\mathbf{e}, a, a^2\}$,

$|St(\pi)| = 1$, $|Z(\pi)| = 3$ and $|\pi| = 4$.

(4) One partition

$\pi = \{\{\mathbf{e}\}, \{a, a^2\}, \{b, ba, ba^2\}\}$,

$St(\pi) = \{\mathbf{e}\}$, $Z(\pi) = \{\mathbf{e}, b, ba, a^2\}$,

$|St(\pi)| = 1$, $|Z(\pi)| = 4$ and $|\pi| = 3$.

(5) Three partitions of the form

$\pi = \{\{\mathbf{e}, a, a^2\}, \{b\}, \{ba, ba^2\}\}$,

$St(\pi) = \{\mathbf{e}\}$, $Z(\pi) = \{\mathbf{e}, a, a^2, b\}$,

$|St(\pi)| = 1$, $|Z(\pi)| = 4$ and $|\pi| = 3$.

(6) Three partitions of the form

$\pi = \{\{\mathbf{e}, b\}, \{a, a^2, ba, ba^2\}\}$,

$St(\pi) = \{\mathbf{e}, b\}$, $Z(\pi) = \{\mathbf{e}, b\}$,

$|St(\pi)| = 2$, $|Z(\pi)| = 2$ and $|\pi| = 2$.

(7) One partition

$\pi = \{\{\mathbf{e}, a, a^2\}, \{b, ba, ba^2\}\}$,

$St(\pi) = \{\mathbf{e}, a, a^2\}$, $Z(\pi) = D_3$,

$|St(\pi)| = 3$, $|Z(\pi)| = 6$ and $|\pi| = 2$.

(8) The coarsest partition

$\pi = D_3$,

$St(\pi) = D_3$, $Z(\pi) = D_3$,

$|St(\pi)| = 6$, $|Z(\pi)| = 6$ and $|\pi| = 1$.

Let Π_{D_3} denote the set of all optimal partitions of D_3, then $|\Pi_{D_3}| = 14$ and let $P = (\Pi_{D_3}, \leq)$ denote the poset of all optimal partitions of D_3. We use the recursive formula (3.5) to calculate the values of the Möbius function $\mu(x, y)$, for $x, y \in P$.

Computing $\mu(1, y)$, for $y = 1, 2, 3, 4, 5$:

(1) 6, 6, 1:

$\mu(1, 1) = 1$ (by the definition of μ)

(2) 2, 2, 2 and 3, 6, 2:

$$\mu(1, 2) = -\sum_{1 \leq z < 2} \mu(1, z) = -\mu(1, 1) = -1$$

(3) 1, 4, 3:

$\mu(1, 3) = -\sum_{1 \leq z < 3} \mu(1, z) = -[\mu(1, 1) + \mu(1, 2)] = 0$ (left of Hasse diagram)

$$\begin{aligned}\mu(1, 3) &= -\sum_{1 \leq z < 3} \mu(1, z) = -[\mu(1, 1) + \mu(1, 2)] \\ &= -[-1 + 1] = 0 \text{ (right of Hasse diagram)}\end{aligned}$$

(4) 1, 2, 4:

$$\begin{aligned}\mu(1, 4) &= -\sum_{1 \leq z < 4} \mu(1, z) = -[\mu(1, 1) + \mu(1, 2) + \mu(1, 3)] \\ &= -[1 - 1 - 1 + 0] = 1\end{aligned}$$

(5) 1, 3, 4:

$$\begin{aligned}\mu(1, 4) &= -\sum_{1 \leq z < 4} \mu(1, z) = -[\mu(1, 1) + \mu(1, 2) + \mu(1, 3)] \\ &= -[1 - 1 + 0] = 0\end{aligned}$$

(6) 1, 1, 5:

$$\begin{aligned}\mu(1,\,5) &= -\sum_{1\le z<5}\mu(1,\,z) = -[\mu(1,\,1)+\mu(1,\,2)+\mu(1,\,3)+\mu(1,\,4)]\\ &= -[0+0-1+1] = 0.\end{aligned}$$

Lastly we represent P in the form of a Hasse diagram. Please note that this Hasse diagram includes the parameters $|St(\pi)|$, $|Z(\pi)|$ and $|\pi|$, for all $\pi \in \Pi_{D_3}$, in the respective vertices and in that order with the corresponding values of $\mu(1,\,y)$ located below in red.

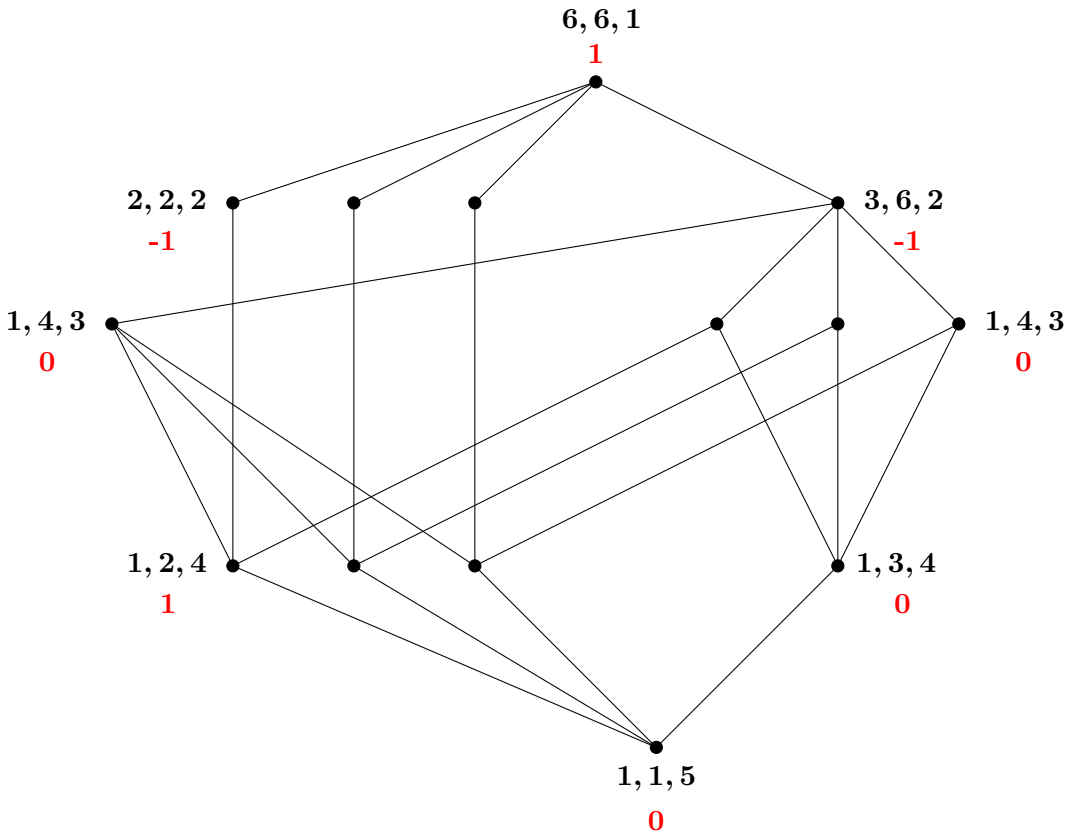

Figure 5.3: Hasse diagram of P.

Substituting these parameters into formulas (5.4) and (5.5), we obtain ex-

pressions for the number of symmetric colorings and symmetric necklaces of D_3.

$$\begin{aligned}
|S_r(D_3)| &= |D_3| \sum_{x\in P} \sum_{y\leq x} \frac{\mu(y,x)}{|Z(y)|} r^{|x|} \\
&= 6\left[r^5 + 3r^4\left(\frac{1}{2}-1\right) + r^4\left(\frac{1}{3}-1\right) + 3r^3\left(\frac{1}{4}-\frac{1}{2}-\frac{1}{3}+1\right)\right. \\
&+ r^3\left(\frac{1}{4}-\frac{3}{2}+2\right) + 3r^2\left(\frac{1}{2}-\frac{1}{2}\right) \\
&\left.+ r^2\left(\frac{1}{6}-\frac{1}{4}-\frac{3}{4}+\frac{3}{2}+\frac{2}{3}-2\right) + r\left(\frac{1}{6}-\frac{1}{6}-\frac{3}{2}+\frac{3}{2}\right)\right] \\
&= 6\left[r^5 - \frac{3}{2}r^4 - \frac{2}{3}r^4 + \frac{5}{4}r^3 + \frac{3}{4}r^3 - \frac{2}{3}r^2 + 0r\right] \\
&= 6r^5 - 13r^4 + 12r^3 - 4r^2
\end{aligned}$$

and

$$\begin{aligned}
|S_r(D_3)/\sim| &= \sum_{x\in P} \sum_{y\leq x} \frac{\mu(y,x)|St(y)|}{|Z(y)|} r^{|x|} \\
&= r^5 + 3r^4\left(\frac{1}{2}-1\right) + r^4\left(\frac{1}{3}-1\right) + 3r^3\left(\frac{1}{4}-\frac{1}{2}-\frac{1}{3}+1\right) \\
&+ r^3\left(\frac{1}{4}-\frac{3}{2}+2\right) + 3r^2\left(\frac{2}{2}-\frac{1}{2}\right) \\
&+ r^2\left(\frac{3}{6}-\frac{1}{4}-\frac{3}{4}+\frac{3}{2}+\frac{2}{3}-2\right) + r\left(\frac{6}{6}-\frac{3}{6}-\frac{6}{2}+\frac{3}{2}\right) \\
&= r^5 - \frac{3}{2}r^4 - \frac{2}{3}r^4 + \frac{5}{4}r^3 + \frac{3}{4}r^3 + \frac{3}{2}r^2 - \frac{1}{3}r^2 - r \\
&= r^5 - \frac{13}{6}r^4 + 2r^3 + \frac{7}{6}r^2 - r.
\end{aligned}$$

Hence we have the following proposition,

Proposition 5.2.1. *For $r \in \mathbb{N}$,*

$$|S_r(D_3)| = 6r^5 - 13r^4 + 12r^3 - 4r^2 \tag{5.10}$$

and

$$|S_r(D_3)/\sim| = r^5 - \frac{13}{6}r^4 + 2r^3 + \frac{7}{6}r^2 - r. \tag{5.11}$$

Example 5.2.1. Let us take $r = 4$. The total number of 4-colorings of D_3 is $4^6 = 4096$. Utilising the above proposition,

$$\begin{aligned} |S_4(D_3)| &= 6(4)^5 - 13(4)^4 + 12(4)^3 - 4(4)^2 \\ &= 3520 \text{ and} \\ |S_4(D_3)/\sim| &= 4^5 - \frac{13}{6}(4)^4 + 2(4)^3 + \frac{7}{6}(4)^2 - 4 \\ &= 612. \end{aligned}$$

5.3 Symmetric Colorings of the Quaternion group Q_8

The total number of r-colorings of Q_8 is r^8. Expressions for the number of symmetric r-colorings of Q_8 are derived in a manner that is precisely identical to the previous section's, hence we skip the procedure, which can be found in [47], and merely state the proposition. Recall that $Q_8 = \{\pm\mathbf{1}, \pm i, \pm j, \pm k\}$.

Proposition 5.3.1. *For $r \in \mathbb{N}$,*

$$|S_r(Q_8)| = 4r^5 - 3r^4 \tag{5.12}$$

and

$$|S_r(Q_8)/\sim| = \frac{1}{2}r^5 - \frac{1}{4}r^4 + \frac{3}{4}r^2. \tag{5.13}$$

Proof. See [47]. □

Example 5.3.1. Let us take $r = 3$. The total number of 3-colorings of Q_8 is $3^8 = 6561$. Utilising the above proposition,

$$\begin{aligned} |S_3(Q_8)| &= 4(3)^5 - 3(3)^4 \\ &= 729 \text{ and} \\ |S_3(Q_8)/\sim| &= \frac{1}{2}(3)^5 - \frac{1}{4}(3)^4 + \frac{3}{4}(3)^2 \\ &= 108. \end{aligned}$$

5.4 Symmetric Colorings of Cyclic groups C_n

It is a far simpler task to calculate the symmetric colorings for cyclic groups C_n of order n, using formulas (5.6) and (5.7). The values $|S_r(C_n)|$ and $|S_r(C_n)/\sim|$, for $n = 1, 2, \cdots, 20$, are summarised in Table 5.1.

C_n	$\lvert S_r(C_n)\rvert$	$\lvert S_r(C_n)/\sim\rvert$
C_1	r	r
C_2	r^2	$\frac{1}{2}(r^2+1)$
C_3	$-2r+3r^2$	r^2
C_4	$-r^2+r^3$	$\frac{1}{2}(r^3+1)$
C_5	$-4r+5r^3$	r^3
C_6	$-2r^2+3r^4$	$\frac{1}{2}(r^4+r^2)$
C_7	$-6r+7r^4$	r^4
C_8	$-r^2-2r^3+4r^5$	$\frac{1}{2}(r^5+r)$
C_9	$-2r-6r^2+9r^5$	r^5
C_{10}	$-4r^2+5r^6$	$\frac{1}{2}(r^6+r^3)$
C_{11}	$-10r+11r^6$	r^6
C_{12}	$2r^2-4r^3-3r^4+6r^7$	$\frac{1}{2}(r^7+r^2)$
C_{13}	$-12r+13r^7$	r^7
C_{14}	$-6r^2+7r^8$	$\frac{1}{2}(r^8+r^4)$
C_{15}	$8r-12r^2-10r^3+15r^8$	r^8
C_{16}	$-r^2-2r^3-4r^5+8r^9$	$\frac{1}{2}(r^9+r)$
C_{17}	$-16r+17r^9$	r^9
C_{18}	$-2r^2-6r^4+9r^{10}$	$\frac{1}{2}(r^{10}+r^5)$
C_{19}	$-18r+19r^{10}$	r^{10}
C_{20}	$-5r^2-8r^3-5r^6+10r^{11}$	$\frac{1}{2}(r^{11}+r^3)$

Table 5.1: Symmetric colorings of C_n.

Chapter 6

Symmetric Bracelets

In this chapter we deliberate the issue of counting the number of symmetric bracelets. Technically, a bracelet is a necklace that must be equivalent under rotations and reflections. In other words, one bracelet may be transformed into another by some rotation and reflection, which is the action of the Dihedral group, hence a bracelet may be identified as the orbit of the Dihedral group's action. The material in this chapter corresponds to [44].

Definition 6.0.1. The equivalence class of all r-colorings of the vertices of a regular n-gon, where all rotations and reflections are taken as equivalent, is called an r-ary *bracelet* of length n. The number of r-ary bracelets of length n is denoted by $B_r(n)$.

Theorem 6.0.1 (Bracelet Counting Theorem Version 1)**.** *Let n, $r \in \mathbb{Z}^+$, the number of r-ary bracelets of $\mathbb{Z}_n$ is*

$$B_r(n) = \frac{1}{2}N_r(n) + \begin{cases} \frac{1}{4}(r+1)r^{\frac{n}{2}}, & \text{for } n \text{ even}, \\ \frac{1}{2}r^{\frac{n+1}{2}}, & \text{for } n \text{ odd}. \end{cases} \tag{6.1}$$

Proof. Refer to [44]. □

The distinguishing factor between a necklace and a bracelet is that bracelets must also be equivalent under reflections. Then the number of *symmetric r-ary bracelets* of $\mathbb{Z}_n$, denoted by $B_r^*(n)$, ends up being the same as the symmetric necklaces, $N_r^*(n)$.

Theorem 6.0.2 (Bracelet Counting Theorem Version 2)**.** *Let n, $r \in \mathbb{Z}^+$. The number of symmetric r-ary bracelets of $\mathbb{Z}_n$ is*

$$B_r^*(n) = N_r^*(n) = \begin{cases} \frac{1}{2}r^{\frac{n}{2}}(r+1), & \text{for } n \text{ even}, \\ r^{\frac{n+1}{2}}, & \text{for } n \text{ odd}. \end{cases} \tag{6.2}$$

Proof. Refer to [44]. □

Let A be a finite abelian group and let $C_2 = \{\mathbf{e}, c : c^2 = \mathbf{e}\}$ and take $G = Dih(A) = C_2 \ltimes_\phi A$. Then $[G : A] = 2$, $G/A = Ac$ and $ca = a^{-1}c\,, \forall a \in A$. Now there exists a (right) group action of G on A defined as follows;

$$ax = a \cdot x \;\text{ and }\; (ac)x = a \cdot x^{-1}.$$

The map $f : A \longrightarrow A$, $f(a) = ax^{-1}$ is called the symmetry on A.

Let $\chi \in r^A$, then $\chi : A \longrightarrow [r]$ and the action of G on A invokes an action on r^A defined as follows;

$$g\chi(x) = \chi(g^{-1}x),\ g \in G \ \text{ and } \ a \in A.$$

Now consider $a \in A$, then

$$a\chi(x) = \chi(xa^{-1}) \ \text{ and } \ (ac)\chi(x) = \chi(xa^{-1}).$$

The map $h : \chi \longrightarrow (ac)\chi$ where $(ac)\chi \in r^A$, is called the symmetry on r^A.

Definition 6.0.2. Let $\chi \in r^A$, then χ is called symmetric if there exists $a \in A$ such that $\chi(x) = \chi(ax^{-1}), \forall x \in A$.

We may now introduce more succinct definitions of necklaces and bracelets using the above constructions.

Definition 6.0.3. Let $\chi \in r^A$, $A\chi$ is called the A-orbit of χ and is known as the r-ary necklace on A. Similarly $G\chi$, the G-orbit of χ, is the r-ary bracelet on A.

Definition 6.0.4. The necklace $A\chi$ is called *symmetric* if there exists a symmetric $\gamma \in A\chi$. Similarly the bracelet $G\chi$ is symmetric if there exists a symmetric $\delta \in G\chi$.

Theorem 6.0.3. *Let A be a finite abelian group and $r \in \mathbb{N}$,*

$$B_r(A) = \frac{1}{2}N_r(A) + \frac{r^{\frac{|A|}{2}}}{2|A[2]|}\left(r^{\frac{|A[2]|}{2}} + |A[2]| - 1\right). \tag{6.3}$$

Proof. Refer to [44]. □

Proposition 6.0.4. *For $n \in \mathbb{Z}^+$,*

$$\mathbb{Z}_n[2] = \begin{cases} 2, & \text{for } n \text{ even} \\ 1, & \text{for } n \text{ odd.} \end{cases}$$

Proof. Recall that the Boolean group

$$\begin{aligned} \mathbb{Z}_n[2] &= \{x \in \mathbb{Z}_n \,:\, 2x \equiv 0 \,(\mathrm{mod}\, n)\} \\ &= \begin{cases} 2, & \text{for } n \text{ even} \\ 1, & \text{for } n \text{ odd.} \end{cases} \end{aligned}$$

□

In a special case of formula (6.3), we take $A = \mathbb{Z}_n$ and combine it with the above proposition to obtain the standard formula for counting bracelets, (6.1).

Lemma 6.0.5. *Let $\chi \in r^A$, then the following statements are equivalent;*

(1) *χ is symmetric,*

(2) *every $t \in A\chi$ is symmetric,*

(3) *$A\chi$ is symmetric,*

(4) *$G\chi$ is symmetric,*

(5) *$G\chi = A\chi$ and*

(6) *$c\chi = A\chi$.*

Proof. **(1)$\Longrightarrow$(2)** Choose x, $y \in A$ such that $xc\chi = \chi$ and $y\chi = t$. Therefore

$$xy^2ct = xy^2cy\chi = xyc\chi = yxc\chi = y\chi = t.$$

(2)$\Longrightarrow$(3) and **(3)$\Longrightarrow$(4)** are both trivial.

(4)$\Longrightarrow$(5) Choose $h \in G\chi$ and $x \in A$ such that $xch = h$. Therefore $\forall y \in A$,

$$ych = ycxch = yx^{-1}h.$$

Hence $Gh = Ah$ and since $A\chi \subseteq G\chi = Gh = Ah \implies A\chi = Ah = G\chi$.

(5)$\Longrightarrow$(6) Trivial.

(6)$\Longrightarrow$(1) Choose $x \in A$ such that $c\chi = x\chi \Longrightarrow x^{-1}c\chi = \chi$. □

Lemma 6.0.6. *Let $\chi \in r^A$ and suppose $G\chi$ is not symmetric. Then $G\chi = A\chi \cup Ac\chi$ where $A\chi \cap Ac\chi = \emptyset$.*

Proof. Lemma 6.0.5 shows that $A\chi$ and $Ac\chi$ are both distinct. It is obvious that $A\chi \cup Ac\chi \subseteq G\chi$ and since $(xc)\chi = x(c\chi) \implies G\chi \subseteq A\chi \cup Ac\chi$, hence we have that $G\chi = A\chi \cup Ac\chi$. □

Proposition 6.0.7. *For $r \in \mathbb{N}$ and a finite abelian group A,*

$$B_r^*(A) = N_r^*(A) = 2B_r(A) - N_r(A)$$

Proof. We may deduce from Lemma 6.0.5 that an r-ary bracelet (G-orbit) is symmetric $\iff$ it is a symmetric r-ary necklace (A-orbit). Specifically $B_r^*(A) = N_r^*(A)$ and by Lemma 6.0.6, if an r-ary bracelet (G-orbit) is not symmetric, then it is equal to the disjoint union of two r-ary necklaces (A-

orbits). Hence

$$\begin{aligned} B_r(A) &= N_r^*(A) + \frac{1}{2}\left[N_r(A) - N_r^*(A)\right] \\ &= \frac{1}{2}\left[N_r(A) + N_r^*(A)\right]. \end{aligned}$$

Therefore

$$N_r^*(A) = 2B_r(A) - N_r(A).$$

□

Hence the last fundamental result is achieved.

Theorem 6.0.8. *For $r \in \mathbb{N}$ and a finite abelian group A,*

$$B_r^*(A) = \frac{r^{\frac{|A|}{2}}}{|A[2]|}\left(r^{\frac{|A[2]|}{2}} + |A[2]| - 1\right). \tag{6.4}$$

Proof. This follows immediately from Proposition 6.0.7 and Theorem 6.0.3.

□

Substitute $A = \mathbb{Z}_n$ in (6.4) and we receive the canonical formula for counting symmetric bracelets, (6.2).

Example 6.0.1. Let us consider V_4. Now $V_4[2] = V_4$ therefore

$$B_r^*(V_4) = \frac{1}{4}r^2(r^2 + 3).$$

This is identical to $|S_r(V_4)/\sim| = N_r^*(V_4)$ from (5.9).

Chapter 7

Conclusion

This final chapter offers a reflection on the body of results that were covered in this dissertation. The vital theorems will be implicitly restated and placed into context within the overall content. Thereafter we seek to overview the current research and open problems in these fields.

The objectives raised in this thesis were to investigate the symmetries on groups and their relevance to combinatorics, Ramsey theory and symmetric colorings in particular. Chapter 2 outlined several principle results from group theory, which formed the backbone of our work in symmetric colorings. The focal points of this chapter include Boolean groups, group actions, the (generalised) Dihedral group and our first significant result, Burnside's lemma, Lemma 2.5.2. These topics then played a crucial role in symmetric colorings.

Chapter 3 aimed to develop some basic combinatorial concepts, an emphasis

was placed on partially ordered set theory and illustrating posets in Hasse diagram form. Constructing the poset of optimal partitions, Example 3.3.3, was another indispensible result to symmetric colorings. Lastly we integrated poset theory into a discussion on the Möbius function, the highlight of this was the Möbius inversion formula, Theorem 3.4.3.

Chapter 4 tackled, in its entirety, Ramsey theory. A cohesive analysis of the classical results that constitute this field was provided. We first detracted to a preliminary section, which offered some basic graph theory and terminology. This section ended with the combinatorial compactness principle, Theorem 4.1.3, which was instrumental in proving the finite version of Ramsey's theorem for sets.

Several versions of Ramsey's seminal theorem were examined, first in a graph theoretical context and then in terms of sets. Ramsey's theorem for two colors, Theorem 4.2.1, was proven first and then its generalisation to a finite number of colors, Theorem 4.2.2, after which an infinite version of this, Theorem 4.2.3 was given. Hence our discourse on graph Ramsey theory ended and we proceeded to state and prove the infinite Ramsey theorem for sets, Theorem 4.2.4, the finite version of the same theorem, Theorem 4.2.5, follows immediately from the infinite version when combined with the aforementioned compactness principle.

The difficulty of calculating Ramsey numbers was subsequently discussed and all values and bounds known to date were tabulated in Table 4.1. Van

der Waerden's theorem was the next critical result placed under scrutiny. It was stated and proved using the original double induction argument and afterwards, following the same procedure as for Ramsey numbers, all known values and bounds for van der Waerden numbers were also tabulated in Table 4.2. We presented two famous results attributed to Shelah and Gowers, which resolve upper bounds for van der Waerden numbers, Theorems 4.5.1 and 4.5.2 respectively, and Berlekamp's lower bound, Theorem 4.5.3.

Schur's theorem was the next traditional result portrayed. The basic form, Theorem 4.6.1 and then the generalised form, Theorem 4.6.2, was proven as well as another table supplied, Table 4.3, of known Schur numbers after which we advanced to a more superior generalisation of Schur's theorem, Rado's single equation theorem, Theorem 4.7.1. This expositional chapter concluded with the statement of the Hales-Jewett theorem, Theorem 4.8.1, whose proof was too involved for the scope of this dissertation and therefore omitted, we also showed that the Hales-Jewett theorem implies van der Waerdens theorem.

A thorough investigation of symmetric colorings was given in Chapter 5. Classical formulas for counting the number of colorings and necklaces on a finite group G were provided, namely Proposition 5.0.2, Theorem 5.0.3 and Theorem 5.0.7. These formulas were then refined and specialised to count the number of symmetric colorings and equivalence classes of symmetric colorings (symmetric necklaces). Theorem 5.0.5 counted the number of symmetric colorings and symmetric necklaces for abelian groups, then Theo-

rem 5.0.6 generalised this theorem to include all arbitrary groups and lastly Theorem 5.0.8 specified only cyclic groups.

Thereafter we calculated the symmetric colorings of the groups V_4, D_3 and Q_8 in order to demonstrate how formulas (5.2) through (5.7) are used. The lattice of subgroups for V_4 and D_3 were subsequently drawn in Figures 5.1 and 5.2 as well as the Hasse diagram for D_3, Figure 5.3, to help visualise the calculations. This section was finalised with Table 5.1, which presented the symmetric colorings of cyclic groups up to order 20, using formulas (5.6) and (5.7).

Chapter 6 ended our treatment of symmetric colorings with symmetric bracelets. This chapter proceeded in a similar fashion to the previous one by first giving standard formulas, Theorems 6.0.1 and 6.0.2, for counting the number of bracelets for the group $\mathbb{Z}_n$ and then generalising them in Theorems 6.0.3 and 6.0.8 respectively to include all finite abelian groups.

Colorings is a rich and diverse field in mathematics of which there are many unsolved problems, [36] chronicles the cutting-edge research done in this area and the key discoveries that accompany them. Symmetric colorings is widely regarded as a separate branch of Ramsey theory, whose sphere of influence extends throughout numerous mathematical disciplines. The principal paper [7], delivers a survey of results and open problems on symmetric colorings of algebraic and geometric objects and expounds the general layout of open problems in this field. This survey is updated in [8], which remains the most

significant and available source of unanswered questions in colorings.

Bibliography

[1] W. Ackermann, *Zum Hilbertschen Aufbau der reellen Zahlen*, Mathematische Annalen 99, doi:10.1007/BF01459088 (1928), 118–133.

[2] T. Ahmed, *List of all known van Der Waerden numbers*, Concordia University, Montreal, http://users.encs.concordia.ca/ ta_ahmed/vdw.html, (2012).

[3] T. Ahmed, D J. Schaal, *On Generalized Schur Numbers*, Experimental Mathematics 25(2) (2016), 213–218.

[4] M. Aigner, Combinatorial theory, Springer-Verlag, New York-Berlin-Heidelberg, 1979.

[5] M. Aigner, A Course in Enumeration, Springer-Verlag, New York-Berlin-Heidelberg, 2007.

[6] R. B. J. T. Allenby and A. Slomson, How to Count: An Introduction to Combinatorics 2nd ed, Chapman & Hall/CRC Press, Taylor & Francis Group, Boca Raton London, New York, 2011.

[7] T. Banakh, I. Protasov, *Symmetry and Colorings: Some results and open problems*, Gomel University, Voprosy Algebry 17 (2001), 4–15.

[8] T. Banakh, *Symmetry and Colorings: Some results and open problems II*, arXiv:11111.1015, preprint.

[9] E. A. Bender, J. A. Goldman, *On the applications of Möbius inversion in combinatorial analysis*, The American Mathematical Monthly 82(8) (1975), 789–803.

[10] E. R. Berlekamp, *A Construction for Partitions Which Avoid Long Arithmetic Progressions*, Canad. Math. Bull. 11 (1968), 409–414.

[11] W. Burnside, Theory of Groups of Finite Order, Cambridge University Press Warehouse, C. J. Clay & Sons, 1897.

[12] P. J. Cameron, Combinatorics: Topics, Techniques, Algorithms, Cambridge University Press, 1994.

[13] J. Chappelon, M. P. R. Marchena, M. I. S. Domínguez, *Modular Schur numbers*, arXiv:1306.5635, preprint.

[14] D. Dummit and R. Foote, Abstract Algebra 3rd ed, John Wiley & Sons, New York, NY, 2004.

[15] W. T. Gowers, *A new proof of Szemerédi's theorem for arithmetic progressions of length four*, Geom. Funct. Anal. 8 (1998), 529–551.

[16] R. Graham, Rudiments of Ramsey theory, CBMS Regional conference series in mathematics volume 123, American Mathematical Society, 2015.

[17] R. Graham, B. Rothschild and J. Spencer, Ramsey Theory 2nd ed, John Wiley & Sons, 1990.

[18] R. E. Greenwood, A. M. Gleason, *Combinatorial Relations and Chromatic Graphs*, Canad. J. Math. 7 (1955), 1–7.

[19] Yu. Gryshko, *Symmetric colorings of regular polygons*, Ars Combinatoria 78 (2006), 277–281.

[20] A. W. Hales, R. I. Jewett, *Regularity and Positional games*, Trans. Amer. Math Soc. 106 (1963), 222–229

[21] P. R. Herwig, M. J. H. Meule, P. M. van Lambalgen, H. van Maaren, *A New Method to Construct Lower Bounds for van Der Waerden Numbers*, Electron. J. Combinatorics 14 (2007), R6.

[22] D. Hilbert, *Ueber die Irreducibilität ganzer rationaler Functionen mit ganzzahligen Coefficienten*, J. Reine Angew. Math. 110 (1892), 104–129.

[23] M. I. Kargapolov and Ju.I. Merzljakov, Fundamentals of the Theory of Groups, Springer-Verlag, New York-Berlin-Heidelberg, 1979.

[24] I. Kashuba, Yu. Zelenyuk, *The number of symmetric colorings of the dihedral group D_3*, Quaestiones Mathematicae, 39:1, DOI: 10.2989/16073606.2015.1015646 (2016), 65–71.

[25] B. M. Landman and A. Robertson, Ramsey theory on the Integers, Student Mathematical Library volume 24, American Mathematical Society, 2004.

[26] B. M. Landman, A. Robertson, C. Culver, *Some New Exact van der Waerden numbers*, Integers: Electron. J. Combinatorial Number Theory 5(2) (2005), A10.

[27] P. A. MacMahon, *Application of a theory of permutations in circular procession to the theory of numbers*, Proc. London Math. Soc. 23 (1892), 305–313.

[28] D. R. Mazur, Combinatorics A Guided Tour, The Mathematical Association of America, MAA Textbooks, 2010.

[29] R. Rado, *Studien zur Kombinatorik*, Math. Zeitschrift. 36 (1933), 424–480.

[30] S. P. Radziszowski, *Small Ramsey Numbers*, Electron. J. Combinatorics, Dynamic Surveys 1 (2006), DS1.

[31] F. P. Ramsey, *On a problem of formal logic*, Proc. Lond. Math. Soc. Ser. 2 30(4) (1930), 338–384.

[32] J. Rotman, Advanced Modern Algebra 1st ed, Prentice Hall, 2002.

[33] J. Rotman, An Introduction to the Theory of Groups, Springer-Verlag, New York, 1994.

[34] I. Schur, *Über die Kongruenz*, $x^m + y^m \equiv z^m \,(mod\, p)$, Jahresbericht der Deutschen Mathematiker-Vereinigung 25 (1916), 114–117.

[35] S Shelah, *Primitive recursive bounds for van der Waerden numbers*, J. Amer. Math. Soc. 1 (1988), 635–636.

[36] A. Soifer, The Mathematical Coloring Book: Mathematics of Coloring and the Colorful Life of Its Creators, Springer, New York, 2009.

[37] A. Soifer, Ramsey Theory: Yesterday, Today, and Tomorrow, Progress in Mathematics volume 285, 2011.

[38] R. P. Stanley, Enumerative Combinatorics 2nd ed volume 1, Cambridge University Press, 2012.

[39] M. Suzuki, *On the Lattice of Subgroups of Finite Groups*, Trans. Amer. Math. Soc. 70(2) (1951), 345–371.

[40] G. E. W. Taylor, Ramsey Theory, Dissertation University of Birmingham, 2006.

[41] B. L. van der Waerden, *Beweis einer audetschen Vermutung*, Nieuw Archief voor Wiskunde 15 (1927), 212–216.

[42] M. Walters, *Combinatorial proofs of the polynomial van der Waerden theorem and the polynomial Hales-Jewett theorem*, J. Lon. Math. Soc. 61 (2000), 1–12.

[43] Yu. Zelenyuk, *Counting symmetric colorings of the vertices of a regular polygon*, Bull. Aust. Math. Soc. 90 (2014), 1–8.

[44] Ye. Zelenyuk, Yu. Zelenyuk, *Counting symmetric bracelets*, Bull. Aust. Math. Soc. 8 (2014), 431–436.

[45] Yu. Zelenyuk, *Monochrome symmetric subsets in colorings of finite abelian groups*, Symmetry 3 (2011), 126–133.

[46] Yu. Zelenyuk, *Symmetric colorings of finite groups*, Groups St Andrews in Bath, LMS Lecture Note Series (2011), 580–590.

[47] Yu. Zelenyuk, *The number of symmetric colorings of the Quaternion Group*, Symmetry 2 (2010), 69–75.